



KONINKRIJK BELGIË

Federale Overheidsdienst

Buitenlandse Zaken,

Buitenlandse Handel en

Ontwikkelingssamenwerking

December 2017

STUDIE OVER DE UITWERKING VAN EEN INSTRUMENT VOOR DE CERTIFICERING VAN DE EVALUATIESYSTEMEN VAN DE ACTOREN VAN DE BELGISCHE ONTWIKKELINGSSAMENWERKING



Federale Overheidsdienst Buitenlandse Zaken, Buitenlandse Handel en
Ontwikkelingssamenwerking

**Dienst Bijzondere Evaluatie van de Belgische
Ontwikkelingssamenwerking**

Studie over de uitwerking van een instrument voor de certificering van de evaluatiesystemen van de actoren van de Belgische ontwikkelingssamenwerking



Corina Dhaene (ACE Europe) en Franck Signoret (COTA)

**Finaal rapport
December 2017**

De studie werd uitgevoerd door een consortium van drie organisaties, met name ACE Europe, COTA en HIVA met de steun van een begeleidingscomité. De Dienst Bijzondere Evaluatie verzekerde de conformiteit van dit evaluatieverslag met de referentietermen.

De meningen die in dit document worden geformuleerd, vertegenwoordigen de standpunten van de auteurs en weerspiegelen niet noodzakelijk de positie van de FOD Buitenlandse Zaken, Buitenlandse Handel en Ontwikkelingssamenwerking.

© FOD Buitenlandse Zaken, Buitenlandse Handel en Ontwikkelingssamenwerking

December 2017

Drukwerk: Drukkerij FOD

Evaluatie nr. S4/2017/01

Wettelijk depot: 0218/2018/007

Dit document is tevens beschikbaar in pdf-formaat in Nederlands en Frans op de website [https://diplomatie.belgium.be/nl/Beleid/Ontwikkelingssamenwerking/Onze werkmethode/Dienst Bijzondere Evaluatie/](https://diplomatie.belgium.be/nl/Beleid/Ontwikkelingssamenwerking/Onze_werkmethod en/Dienst_Bijzondere_Evaluatie/) of bij de Dienst Bijzondere Evaluatie.

Het verslag wordt als volgt geciteerd:

DBE 2017, *Studie over de uitwerking van een instrument voor de certificering van evaluatiesystemen van actoren van de Belgische ontwikkelingssamenwerking*, FOD Buitenlandse Zaken, Buitenlandse Handel en Ontwikkelingssamenwerking, Brussel.

Woord vooraf

Het consortium kreeg in het voorjaar 2017 het mandaat van de Dienst voor Bijzondere Evaluatie (DBE) van de Belgische Ontwikkelingssamenwerking van de Federale Overheidsdienst Buitenlandse Zaken, Buitenlandse Handel en Ontwikkelingssamenwerking om na te denken over de definitie van een mechanisme voor certificering van de evaluatiesystemen van de actoren, gefinancierd door de Belgische ontwikkelingssamenwerking.

De certificering is in de eerste plaats een instrument waarvan de doelstelling helder en lovenswaardig is : bevorderen van de kwaliteit van de evaluaties van de actoren die gefinancierd worden met Belgische fondsen en bijdragen tot de capaciteit van de actoren om te leren van wat werd gerealiseerd op het terrein en op gepaste wijze verantwoording af te leggen.

Het consortium heeft op deze doelstelling verder gewerkt, met het oog op het ondersteunen van leren en een continue verbetering van de praktijken inzake evaluatie.

Hiervoor ontwikkelde het consortium een proces dat lang leek voor sommigen en misschien te kort voor anderen. Het belangrijkste is dat het consortium zich ingespannen heeft om de geïnteresseerde stakeholders te verzamelen rond het denkproces. De actoren werden zowel indirect betrokken (bijv. door de deelname van de Federaties in het begeleidingscomité) als direct (bijv. door focusgroep discussies, sense-making workshops en individuele interviews).

Het was onze wens om te komen tot een instrument dat rekening hield met de belangen van de diverse stakeholders (Minister, de DBE, de Directie-Generaal Ontwikkelingssamenwerking en Humanitaire Hulp (DGD), BTC, BIO, de actoren van de niet-gouvernementele samenwerking (ANGS) en de Institutionele Actoren (IA)) met de wet als uitgangspunt.

We hebben een instrument ontwikkeld en getest dat, volgens ons, iedere organisatie de kans geeft om haar praktijken inzake evaluatie te verbeteren met als voornaamste doelstelling : zichzelf en haar praktijk geregeld in vraag te stellen zodat ze haar benadering, methoden en interventiestrategieën kan aanpassen en de kwaliteit van haar interventies kan verbeteren.

Het proces van deze studie heeft veel organisaties gemobiliseerd en we beseffen dat hun tijd kostbaar is. Om die reden willen we alle mensen en organisaties die werden gemobiliseerd en die hun steentje hebben bijgedragen danken. We vermelden in het bijzonder de constructieve inzet van de drie Federaties FIABEL, ACODEV en NGO-Federatie.

Deze studie was, tenslotte, niet tot stand kunnen komen zonder de niet aflatende inzet en actieve deelname van de DBE. Het werk vond aldus plaats in een sfeer van vertrouwen. We danken hiervoor de Dienst en het voltallige personeel.

Het consortium ACE Europe, COTA, HIVA KULeuven

Corina Dhaene, Franck Signoret, Bruno Mola, Heinz Werner Engel, Geert Phlix, Gerda Heyde en Jan Van Ongevalle.

Inhoudstafel

Woord vooraf.....	3
Inhoudstafel.....	5
Afkortingen en acroniemen	6
Beknopte samenvatting.....	7
1. Inleiding	11
2. Bevindingen van de eerste fase en conclusies voor de opzet van het instrument	15
2.1. Bevindingen over de capaciteit van de actoren	15
2.1.1. Capaciteit van actoren: ANGS, BTC en BIO	15
2.1.2. Capaciteit van de andere stakeholders	24
2.1.3. Informatienoden van de actoren en stakeholders.....	25
2.2. Bevindingen vanuit de literatuur	26
2.3. Conclusies over de opzet en uitwerking van het instrument	29
3. Bevindingen van de testfase	32
3.1. Uitvoering van de testfase	32
3.2. Conclusies over de elementen die getest moesten worden	33
3.3. Gevolgen voor de uitwerking van het instrument.....	34
4. Beschrijving van het instrument	37
4.1. Situering van het instrument	37
4.2. Doelstelling en scope van het instrument.....	40
4.3. Hoe zal het instrument werken?	42
4.4. Modaliteiten voor het beheer	46
5. Overzicht van de bijbehorende documenten	49
5.1. Overzicht van de criteria.....	49
5.2. Handleiding voor zelfevaluatie	51
5.3. Handleiding voor de assessor	51
Bijlage A: Overzicht van documenten die werden geconsulteerd.....	53
Bijlage B: Overzicht van organisaties die werden geconsulteerd	57
Bijlage C: Overzicht van key-events	58
Bijlage D: Overzicht van organisaties uit de testfase.....	59

Afkortingen en acroniemen

ANGS	Actoren van de Niet-Gouvernementele Samenwerking	
BIO	Belgische Investeringsmaatschappij voor Ontwikkelingslanden	
BTC	Belgisch Technische Coöperatie	
DAC	Development Assistance Committee	
DBE	Dienst Bijzondere Evaluatie	
DGD	Directie-Generaal Ontwikkelingssamenwerking en Humanitaire Hulp	
DFI	Development Finance Institute	
ECB	Evaluation Capacity Building	
EFQM	European Foundation for Quality Management	
FTE	Full Time Equivalent	
GSK	Gemeenschappelijk Strategisch Kader	
IA	Institutionele Actoren	
ITAD	Information Technology and Agriculture Development	
LEADS	L = Little or no elements can be identified; E = Elements do exist; concrete Action is taken; D = largely Developed; S = Sustainable	A =
M&E	Monitoring en Evaluatie	
MEL	Monitoring Evaluation Learning	
NGO	Niet-Gouvernementele Organisatie	
NGSOC	Overlegstructuur tussen de ANGS en de subsidiërende overheid	
OESO	Organisatie voor Economische Samenwerking en Ontwikkeling	
SDG	Sustainable Development Goal	
SMART	Specific, Measurable, Achivable, Relevant, Time-oriented	
ToC	Theory of Change	
VLIR-UOS	Vlaamse interuniversitaire Raad – Universitaire Ontwikkelingssamenwerking	

Beknopte samenvatting

Dit rapport beschrijft de resultaten van een participatief ontwerpproces van een instrument voor de certificering van de evaluatiesystemen van Belgische ontwikkelingssamenwerkingsactoren (88 in totaal).

Doelstelling en rationale van het certificerings-instrument

Het regelgevend kader voor de Belgische ontwikkelingssamenwerking bepaalt dat de actoren van de niet-gouvernementele samenwerking (ANGS) en de institutionele actoren (IA) die recht hebben op overheidssubsidies alsook de Belgische Technische Coöperatie (BTC) en de Belgische Investeringsmaatschappij voor Ontwikkelingslanden (BIO) verantwoordelijk zijn voor de evaluatie van hun interventies. Om de kwaliteit van deze evaluaties te garanderen werd de Dienst Bijzondere Evaluatie (DBE) belast met de certificering van de evaluatiesystemen van de actoren en het ex-post evalueren van de kwaliteit van deze systemen. Hiertoe werd een instrument uitgewerkt dat enerzijds toelaat om de kwaliteit van de evaluatiesystemen vast te stellen en anderzijds instrumenten aanbiedt om deze evaluatiesystemen op een continue manier te versterken. Het instrument kreeg de naam Evaluation Building Capacity of ECB. Het instrument leidt tot het verlenen van een certificaat voor Evaluation Capacity Building of ECB-certificaat. Het certificaat bevestigt de kwaliteit van het (externe) evaluatiesysteem, de evaluatiepraktijk van een organisatie en haar engagement tot verbetering.

De focus van het instrument is het evaluatiesysteem van de actoren. Wat het opvolgingssysteem (monitoring) betreft, houdt het instrument rekening met de capaciteit om betrouwbare, bruikbare en relevante gegevens te verschaffen, waar de evaluaties op kunnen bouwen.

Participatieve design- en testfase

Het certificeringsinstrument werd ontwikkeld in nauw overleg met alle betrokken actoren waaronder ook de organisaties die gecertificeerd willen worden en vertegenwoordigers van DGD en de strategische cel van de Minister voor Ontwikkelingssamenwerking. Deze actoren waren ook vertegenwoordigd in het begeleidingscomité dat de DBE adviseerde over de uitvoering en het resultaat van de studie. De DBE was actief betrokken in de uitvoering van de studie als eigenaar van het instrument.

Het sterk participatieve karakter van de design- en testfase was belangrijk om de relevantie en een brede toepasbaarheid van het instrument te verzekeren voor de ontwikkelingsactoren die sterk verschillen qua omvang, evaluatiecapaciteit en mandaat.

- Tijdens de designfase werd in eerste instantie de huidige capaciteit van de ANGS, BTC en BIO voor het ontwerpen, uitvoeren en gebruiken van evaluaties in kaart gebracht. Dit gebeurde op basis van focusgroepen, interviews en een analyse van relevante studies. Hieruit bleek duidelijk dat alle actoren evaluatie belangrijk vinden en hierin willen investeren. Tegelijk was er consensus over de elementen en stappen die noodzakelijk zijn om te komen tot goede evaluaties. Bovendien werden een aantal

aandachtspunten geïdentificeerd waar organisaties in de toekomst nog dienen op te werken om hun evaluatiecapaciteit te versterken. Op basis van deze inzichten werden de beoordelingscriteria van het certificeringsinstrument uitgewerkt alsook de modaliteiten voor het certificeringsproces.

- Opeenvolgende voorstellen werden steeds voorgelegd aan de sector door middel van sense-making workshops en verschillende bijeenkomsten van het begeleidingscomité. Dit was een iteratief proces waarbij voorstellen vaak werden herwerkt en verder gepreciseerd naargelang de vooruitgang in de inzichten/kennissen.
- De testfase, waaraan 8 organisaties deelnamen, diende vooral om de helderheid en gebruiksvriendelijkheid van het instrument en bijbehorende handleidingen en formats te testen alsook de doeltreffendheid ervan. Tevens werden de rol en opdracht van een externe assessor¹ getoetst. Een belangrijk resultaat van de testfase was de bevestiging dat de basisopzet van het instrument werkt en beantwoordt aan de vooropgezette doelstellingen. Ook werd duidelijk dat de assessoren rekening moeten houden met het spanningsveld tussen 'het ondersteunen van leren' en 'het ter verantwoording roepen' van de organisatie hetgeen inherent bleek aan de kwalitatieve opzet van het instrument. Voorts bevestigde de testfase dat het instrument werkt voor een diversiteit aan organisaties. Dankzij de zelfevaluatiecomponent van het instrument kunnen de organisaties de evaluatiefunctie binnen hun eigen context situeren, waarmee de externe assessor dan rekening kan houden.

Opbouw van het certificeringsinstrument

Het certificeringsinstrument volgt de logica van capaciteitsversterking van evaluatie (Evaluation Capacity Building (ECB)) die gericht is op het duurzaam inbedden van de evaluatiepraktijk in de organisatie en het versterken van haar capaciteit om aan te sturen op betrouwbare, valide en bruikbare evaluatierapporten. Deze ECB-logica vertaalt zich zowel in de inhoud van het instrument als in de uitvoering en de opvolging ervan. Het instrument omvat de volgende hoofdcomponenten:

- **Beoordelingscriteria:** Het instrument bevat duidelijk omschreven criteria en sub-criteria voor de beoordeling van (i) de evaluatiecapaciteit (incl. visie en oriëntatie, taakverdeling, normatief kader, leiderschap en leervermogen), (ii) het evaluatieproces (incl. ontwerp, uitvoering en gebruik van de resultaten) en (iii) de kwaliteit van de evaluatierapporten (validiteit, betrouwbaarheid en bruikbaarheid). Er is bijzondere aandacht voor de integratie van de evaluatiefunctie in de organisatie en in de programmacyclus. De vragen rond de criteria zijn niet binair geformuleerd zodat deze niet eenvoudigweg met ja of neen kunnen beantwoord worden. De bedoeling is niet om de 'compliance' te testen maar wel de 'performance' van de organisatie te analyseren en te waarderen.
- **Zelfevaluatie:** Een zelfevaluatie door de organisatie op basis van de beoordelingscriteria en het opstellen van een verbeterplan vormen de basis van het instrument. De zelfevaluatie en het opstellen van het verbeterplan moet gebaseerd zijn op een collectieve oefening. De organisatie brengt in haar analyse zelf bewijsmateriaal aan voor haar bevindingen en identificeert goede praktijken. De zelfevaluatie bevat een narratieve beoordeling en een score per sub-criterium volgens een 5-puntenschaal op basis van de LEADS² methode. Een organisatie moet voldoende kunnen bewijzen dat ze beantwoordt aan 4 op de 6 criteria die een minimale basiscapaciteit vertegenwoordigen op het vlak van evaluatie alvorens zij kan

¹ Hier moet de term 'assessor' begrepen worden als 'persoon die een assessment uitvoert'

² L = Little or no elements can be identified; E = Elements do exist; A = concrete Action is taken; D = largely Developed; S = Sustainable

deelnemen aan het proces dat leidt tot een ECB-certificaat. Uit de testfase blijkt dat een organisatie minimaal 8 dagen nodig heeft voor het ganse proces (rekening houdend met de betrokkenheid van verschillende mensen).

- **Externe validatie:** De externe validatie gebeurt door een pool van assessoren onder de leiding van de DBE. Een assessor analyseert in eerste instantie de criteria die betrekking hebben op de methodologische kwaliteit van de evaluatierapporten en besteedt een bijzondere aandacht aan de validatie van de criteria die een minimale basiscapaciteit vertegenwoordigen. Vervolgens beoordeelt de assessor de kwaliteit van de collectieve zelfevaluatie, de kwaliteit van het evaluatiesysteem (in zijn geheel) en de coherentie en kwaliteit van het verbeterplan. Voor deze drie rubrieken geeft de assessor een narratieve beoordeling en een score volgens een 5-puntenschaal op basis van de LEADS-methode. De criteria en sub-criteria van het instrument bieden een leidraad voor de beoordeling maar worden niet elk afzonderlijk beoordeeld en gescoord (met uitzondering van de sub-criteria die betrekking hebben op de evaluatierapporten en de analyse van validiteit, betrouwbaarheid en bruikbaarheid). De assessor vertrekt steeds vanuit de context van de organisatie en haar visie op verandering en probeert van daaruit, en op basis van het beschikbare materiaal, 'critical issues' te definiëren en te bespreken. Om het leren te ondersteunen, geeft de assessor in zijn verslag voldoende toelichting (*appreciative feedback*). Op basis van de aanbeveling die de assessor in zijn verslag formuleert, kan de DBE een ECB-certificaat toekennen.
- **Opvolging:** Het delen van ervaringen en (goede) praktijken rond evaluatie en capaciteitsversterking met de andere actoren van de Belgische ontwikkelingssamenwerking is een belangrijke bijdrage tot de ECB-logica. Het instrument voorziet in een jaarlijkse vrijwillige collectieve opvolging van de verbeterplannen en in de uitwisseling van de ervaringen en evaluatiepraktijken door de organisaties. Dit kan gebeuren onder de vorm van een leerevent waarvoor de DBE uitnodigt (in afstemming met de federaties, de ANGS, BTC en BIO) en waarvan telkens een verslag wordt gemaakt zodat opvolging van veranderingen mogelijk is. Het is eveneens voorzien dat de actoren hun verbeterplan agenderen op de institutionele dialoog met de subsidiërende overheid. Voor BTC en BIO valt de opvolging automatisch onder de Raad van Bestuur, waarin ook de overheid vertegenwoordigd is. Om de opvolging te faciliteren, worden het verslag van de assessor en het verbeterplan van de organisatie gedeeld.
- **Periodiciteit:** Een cyclus van certificering loopt samen met de programmacyclus voor de ANGS en met het beheerscontract voor BTC en BIO.

Het initiatief voor Evaluation Capacity Building wordt beschreven in dit rapport. Het initiatief wordt ondersteund door een lijst met criteria, een handleiding voor zelfevaluatie en een handleiding voor de assessor.

Succesfactoren

Een aantal succesfactoren die bepalend zijn voor de doeltreffendheid van het instrument werden geïdentificeerd:

- De DBE staat garant voor de kwaliteit van de uitvoering van het instrument en wordt daarin ook erkend door de stakeholders.
- Organisaties die proactief de evaluaties in samenwerking met hun partners beheren en zelf goed overwegen waarom en hoe deze evaluaties dienen uitgevoerd te worden, hebben meer kans om bruikbare evaluaties te produceren. Vandaar het belang om de evaluatiefunctie goed te integreren in de programmacyclus.
- Er is een minimum engagement nodig bij alle actoren: in een context van besparingen blijkt het in de praktijk moeilijk om middelen voor evaluatie en voor het ECB-initiatief vrij te maken. Het risico bestaat ook dat organisaties hun verbeterplan beperken tot acties die ze met het huidige personeel kunnen uitvoeren zonder bijkomende middelen. Dit kan een hindernis zijn voor innovatie. Daarom is een proactief aanbod van de federaties belangrijk alsook het jaarlijkse leerevent, georganiseerd door de DBE.
- DGD (D3 Directie Civiele Maatschappij en D4 Directie Organisatiebeheer) ontwikkelt een transparante aanpak om meer systematisch gebruik te maken van evaluatierapporten en van hun recht om de referentietermen van evaluaties bij te sturen.

1. Inleiding

Dit rapport bevat een voorstel voor een certificeringsinstrument dat werd uitgewerkt in opdracht van de Dienst Bijzondere Evaluatie die ook de finale beslissingen neemt over het instrument. De opdracht kadert binnen de Wet betreffende de Belgische Ontwikkelingssamenwerking³ en de regelgeving m.b.t. het mandaat van de Bijzondere Evaluator. Uit dat regelgevend kader volgt dat de actoren van de Belgische ontwikkelingssamenwerking verantwoordelijk zijn voor de evaluatie van hun interventies en dat de Dienst Bijzondere Evaluatie (DBE) belast wordt met het certificeren van de evaluatiesystemen en het ex-post evalueren van de kwaliteit ervan. De opdracht werd uitgevoerd door een consortium van drie organisaties: ACE Europe, COTA en HIVA, met een team dat bestond uit zeven consultants onder leiding van Corina Dhaene (ACE Europe) en Franck Signoret (COTA).⁴

De opdracht werd door de DBE afgebakend naar onderwerp, doelgroep en finaliteit:

- **Onderwerp:** het gaat om het evaluatiesysteem van de actoren. Voor wat opvolging betreft, wordt de capaciteit om betrouwbare, bruikbare en relevante gegevens, waar de evaluaties op kunnen bouwen, te voorzien meegenomen in het instrument.⁵ Verder gaat het om de externe evaluatierapporten van de actoren met specifieke aandacht voor de evaluaties van DGD-programma's. DGD heeft bepaald dat ANGS minimaal een tussentijdse en een eindevaluatie van hun programma moeten uitvoeren en heeft ook bepaald aan welke criteria deze evaluaties moeten voldoen.⁶
- **Doelgroep:** het gaat in totaal om 88 actoren, meer bepaald de ANGS die door de overheid werden geaccrediteerd en recht hebben op overheidssubsidies, BTC en BIO.
- **Finaliteit:** in een kadernota over het certificeringsproces verduidelijkte de DBE haar visie op certificering als een instrument voor een aanpak van continue verbetering waaraan geen sancties verbonden zijn.⁷ Hierin verschilt de aanpak van de screening van de processen van de ANGS die werd doorgevoerd in opdracht van DGD met het oog op de accreditatie van de ANGS en hun financiering. Er werd evenwel aangestuurd op complementariteit met de screening: een aantal criteria van de screening omtrent resultaatsgericht beheer werden meegenomen in de ontwikkeling van het instrument.

Tegelijk was duidelijk dat het regelgevend kader veel openlaat en dat de DBE veel vrijheid krijgt om zelf richting te geven aan de certificering en om, in overleg met de actoren, een proces uit te stippelen dat gericht is op de verbetering van de M&E systemen.

³ Wet van 19/03/2013 betreffende de Belgische Ontwikkelingssamenwerking zoals gewijzigd door de wet van 9/01/2014 en de wet van 16/06/2016.

⁴ De andere leden van het team waren: Geert Phlix (ACE Europe), Bruno Mola (COTA), Jan Van Ongevalle (HIVA), Heinz Werner Engel en Gerda Heyde.

⁵ 'Zo zal het monitoringsysteem enkel onderdeel vormen van de certificering voor wat betreft de capaciteit om nuttige informatie voor de evaluaties te voorzien' (DBE (2017) Discussienota van de DBE over de certificering van de M&E systemen van de Belgische Ontwikkelingsactoren). Dit betekent dat er geen rekening wordt gehouden met de jaarlijkse rapportage (narratief en financieel) van de organisaties.

⁶ KB van 11 september 2016 betreffende de niet-gouvernementele samenwerking, art. 44.

⁷ DBE (2017) *Op.Cit.*

Zoals gevraagd in de referentietermen⁸ werd de opdracht uitgevoerd in nauw overleg met alle betrokken actoren omdat het instrument vorm diende te krijgen via een participatief proces met de actoren die achteraf gecertificeerd zullen worden. Participatie werd tevens belangrijk geacht om ervoor te zorgen dat het instrument een echte meerwaarde kan betekenen voor de actoren. Er werd ook rekening gehouden met het feit dat het instrument toepasbaar moest zijn op ontwikkelingsactoren die sterk verschillen qua omvang, evaluatiecapaciteit en mandaat. Het traject van certificering moest budgettair en qua timing realistisch zijn. Tenslotte moest het consortium een instrument voor zelfevaluatie ontwikkelen dat de actoren in staat stelt om zelf de kwaliteit van hun evaluatiesysteem te toetsen. Daarnaast werd gevraagd op welke manier de opdracht van de DBE inzake het ex-post evalueren van kwaliteit geïntegreerd kan worden in het instrument.

Dit rapport bevat de antwoorden op de vragen die werden gesteld in de referentietermen van de opdracht en die te maken hebben met de criteria, de aansturing van het instrument, de periodiciteit, de toekenning, de relatie met de screening en de middelen voor de uitvoering.⁹

Om te komen tot de antwoorden werden de volgende activiteiten ondernomen:

- Informatiesessie voor alle actoren, georganiseerd door de DBE.
- Documentstudie (zie het overzicht van de geconsulteerde documenten in bijlage);
- Interviews met stakeholders in de startfase. Een overzicht van alle geïnterviewde personen kan in bijlage van dit rapport geraadpleegd worden.
- Werksessies/focusgroepen met de actoren van de Belgische ontwikkelingssamenwerking: er werden 5 werksessies georganiseerd met deelnemers uit meer dan 40 verschillende ANGTS, BTC en BIO. In totaal konden dus meer dan de helft van de actoren van de Belgische Ontwikkelingssamenwerking actief betrokken worden. De focus van deze focusgroepen lag op de invulling van de notie 'kwaliteit' en op het externe evaluatieproces. Daarnaast werd ook discussie gevoerd over een aantal specifieke topics: *governance* van de evaluatiefunctie, evalueerbaarheid en Theory of Change (ToC), evalueerbaarheid en beschikbaarheid van gegevens, betrekking van partners.
- Sense-making workshop: er werden in totaal 2 sense-making workshops georganiseerd met 61 deelnemers.
- Een eerste voorstel voor het design van het instrument werd voorgesteld aan het begeleidingscomité in juli 2017.
- Bijkomende interviews met de diverse stakeholders: om de volgende stap (het verder ontwikkelen van het instrument en de documenten) voor te bereiden werden bijkomende interviews georganiseerd, meer bepaald met de federaties van de ANGTS, BTC, de strategische cel en DGD (D3, Directie Civiele Maatschappij en D4, Directie Organisatiebeheer, meer specifiek D4.2, Kwaliteitsbeheer en Resultaten).
- Gedetailleerde uitwerking van het instrument: tussentijdse nota met de beschrijving van het instrument en de opzet voor zelfevaluatie en beoordeling. Deze documenten werden voorgesteld aan het begeleidingscomité van 8 september 2017.
- Testfase: de testfase werd uitgevoerd tussen 12 september 2017 en 23 oktober 2017. Acht organisaties hebben deelgenomen: BTC, Artsen zonder Vakantie, Avocats sans Frontières, VIA Don Bosco, VLIR-UOS, BIS-MSI (ACLB), Caritas België en Entraide et Fraternité. Deze organisaties hebben zich aangemeld na een open oproep via de

⁸ DBE (februari 2017) Vereenvoudigde onderhandelingsprocedure met bekendmaking voor de uitwerking van een certificeringsinstrument voor de evaluatiesystemen van de actoren van de Belgische ontwikkelingssamenwerking. Bestek nr S4/2017/01

⁹ Op.cit. zie onder de punten B6 en B7 (Nederlandstalig document is niet genummerd).

website, de werksessies en de sense-making workshop en vertegenwoordigen een goede mix van grote en kleine organisaties, complexe en minder complexe, organisaties met en zonder gemeenschappelijk DGD-programma, actoren die in diverse sectoren actief zijn, en actoren met zowel Noord als Zuidwerking.¹⁰ De testfase werd ondersteund door de volgende documenten: een PowerPoint presentatie (die kon gebruikt worden door de organisaties om intern te mobiliseren), een handleiding voor zelfevaluatie en een verbeterplan met rapportageformat in Excel en een format voor het verbeterplan in Word, een handleiding voor de externe validatie met format voor verslag en Excel voor berekenen van gewogen totaal van beoordeling;

- Evaluatie van de testfase met de actoren van de testfase, de federaties en de DBE en voorstelling van deze resultaten op het Begeleidingscomité van 7 november 2017;
- Afronden van de testfase en op punt stellen van het instrument en de bijbehorende documenten;
- Restituties (na aanvaarding van het rapport en bijhorende documenten door de DBE). Deze werden georganiseerd in december 2017 met DGD enerzijds en met de actoren anderzijds (20 februari 2018).

De documentstudie, interviews, werksessies en sense-making workshop hebben vooral bijgedragen tot het design van het instrument, het formuleren van criteria (tijdens een evaluatieproces) en het identificeren van enkele hypothesen en risico's van het toekomstige instrument. De testfase diende vooral om de helderheid en gebruiksvriendelijkheid van het instrument en bijbehorende handleidingen en formats te testen alsook de doeltreffendheid ervan.

De DBE was nauw betrokken bij de verschillende stappen en nam actief deel aan de testfase (als co-assessor tijdens bezoeken aan de testorganisatie, voor de analyse van evaluatierapporten, als co-reader van verslagen van de assessor). Dit maakte het mogelijk om eventuele 'conflicts of interest' uit te sluiten (bijv. als leden van het consortium in het recente verleden betrokken waren geweest in de begeleiding van de testorganisaties of het evalueren van hun programma's). Een uitgebreid begeleidingscomité¹¹ heeft bijgedragen tot inhoudelijke discussie en feedback over het instrument en tot bijsturing waar nodig. Zo werd na de bespreking van een eerste voorstel door het begeleidingscomité meer aandacht gevraagd voor de wijze van beoordeling (en het scoren van criteria) en een beter evenwicht tussen het scoren van capaciteit, de evaluatieprocessen en de kwaliteit van de evaluatierapporten.

De uitvoering van de opdracht werd tijdens de uitvoering aangepast. De DBE ging akkoord met een vraag van het begeleidingscomité van 8 september 2017 om de evaluatierapporten van de organisaties in de testfase individueel en in detail te laten analyseren door de assessoren. Dit was niet expliciet gevraagd in de referentietermen en was evenmin opgenomen in het voorstel van het consortium. Om deze bijkomende taak te vervullen werden extra dagen toegevoegd.

Het spanningsveld tussen 'verantwoording afleggen' en 'leren' dat typisch is in evaluaties en in het evaluatiedomein, heeft zich ook in de uitvoering van deze opdracht (evaluatiesystemen) geuit. Het consortium heeft zo goed mogelijk dit spanningsveld erkend en geïntegreerd in het instrument.

¹⁰ Dit betekent dat in de testfase geen ANGS betrokken was die enkel Noordwerking heeft.

¹¹ Het begeleidingscomité is vijf keer samengekomen tijdens het proces. Het comité was samengesteld uit vertegenwoordigers van BIO, BTC, DGD, het Kabinet voor Ontwikkelingssamenwerking, NGO-Federatie, FIABEL en ACODEV en diverse ANGS (Protos, VLIR-UOS, Vredeseilanden, VVOB, Rode Kruis, Broederlijk Delen, WereldSolidariteit, APEFE).

Het rapport start met een overzicht van de belangrijkste bevindingen van de piloot- en testfase en conclusies ten aanzien van het instrument voor certificering, vervolgens wordt het instrument beschreven en worden de vragen van het bestek beantwoord, en tenslotte wordt een overzicht en korte beschrijving van de bijhorende documenten/tools gegeven.

2. Bevindingen van de eerste fase en conclusies voor de opzet van het instrument

Dit hoofdstuk bestaat uit drie paragrafen: (i) bevindingen over de capaciteit van de actoren, (ii) bevindingen uit de literatuur over evaluatie, (iii) conclusies.

2.1. Bevindingen over de capaciteit van de actoren

De eerste fasen van het onderzoek met de documentstudie, focusgroepen en sense-making ateliers hebben veel duidelijk gemaakt over de huidige capaciteit van de actoren, met name op het vlak van de organisatie van de evaluatiefunctie, de evaluatiepraktijk en het begeleiden van externe evaluaties. Interviews gaven inzicht in de capaciteit van de actoren en stakeholders die al dan niet een rol kunnen spelen in de uitvoering van het instrument. Het eerste punt geeft aan waarover het instrument moet gaan. Het tweede punt maakt duidelijk hoe het instrument in de praktijk doeltreffend gebruikt kan worden.

We gaan eerst in op de capaciteit van de ANGS, BTC en BIO omtrent evaluaties.

2.1.1. Capaciteit van actoren: ANGS, BTC en BIO

De analyse van de capaciteit is gebaseerd op de input van de actoren in 5 focusgroepen (georganiseerd in mei 2017), de resultaten van de screening van de capaciteit van de ANGS (overzicht van algemene scores¹² voor de sectie rond resultaatgericht beheer) en literatuur (voornamelijk de DBE studies rond evalueerbaarheid van 2016 en de meta-evaluatie van de kwaliteit van evaluatierapporten van actoren van de niet-gouvernementele samenwerking van 2013).¹³ Voor de analyse van de capaciteit van BTC en BIO werden bijkomende interviews georganiseerd. De bedoeling hiervan was om na te gaan in hoeverre een instrument geschikt kon zijn voor de ANGS enerzijds en BTC en BIO anderzijds. We gaan hier dieper op in op het einde van dit hoofdstuk.

Het was duidelijk dat alle actoren evaluatie belangrijk vinden en hierin willen investeren. De druk van DGD en andere donoren heeft daartoe bijgedragen maar ook de behoefte om beter te begrijpen wat werkt en niet werkt. De ambities van de actoren ten aanzien van evaluatie worden evenwel sterk beïnvloed door wat DGD met de evaluatierapporten zal doen: meer transparantie over hoe rapporten en conclusies gebruikt kunnen en zullen worden door de administratie is een belangrijke vraag bij de actoren.

¹² De scores per organisatie werden niet gedeeld met het consortium.

¹³ DBE/ADE (2013) *Meta-evaluatie van de programma's van de niet-gouvernementele actoren*, DBE/South Research/IOB (2016) *Geloofwaardig en zinvol evalueren; tussen droom en daad. Een studie van de evalueerbaarheid van interventies ge(co-)financierd door de Belgische Samenwerking*.

Er was een brede consensus onder de actoren over de elementen en stappen die verzekerd moeten zijn om te komen tot goede evaluaties. Deze sluiten overigens goed aan bij wat in de literatuur van belang wordt geacht. Nuances in mening zijn terug te brengen tot het kunnen beschikken in organisaties over personen die Monitoring en Evaluatie (M&E) in hun taakomschrijving hebben. Nadat de actoren collectief de stappen hebben aangeduid, hebben ze zichzelf tijdens de focusgroepen gescoord ten aanzien van hun invulling van kwaliteit in antwoord op de vraag: 'waar staan we in de praktijk?'.

Wat hieronder volgt is een overzicht van de belangrijkste bevindingen uit de focusgroepen. De bevindingen volgen het proces van een evaluatie en de voornaamste stappen ervan: ontwerp, uitvoering en gebruik. Iedere stap wordt aangevuld met een analyse van het consortium. Wat we nu al kunnen stellen is dat de hogere scores van de screening met betrekking tot resultaatsgericht beheer allemaal sterk genuanceerd werden tijdens de discussie binnen de focusgroepen. De bevindingen uit de focusgroepen sluiten met andere woorden beter aan bij de conclusies van de DBE-studies rond evalueerbaarheid en meta-evaluatie, die de praktijk strenger beoordelen.

Stappen die belangrijk zijn in een evaluatieproces – de mening van de ANGS	
Ontwerp van een evaluatie	
Goed de scope van de evaluatie bepalen	Voor al actoren met M&E personeel stellen dat dit al bij de start van het programma duidelijk moet zijn. In één focusgroep oordeelden de deelnemers dat ook evaluatievragen best geformuleerd worden bij de start van het programma en dat ze bij het bepalen van de referentietermen verfijnd of opnieuw bekeken kunnen worden. Slechts enkele organisaties verwijzen naar het gebruik van de Theory of Change (ToC) om na te denken over evaluaties. Het formuleren van goede evaluatievragen blijkt moeilijk te zijn voor de meeste deelnemers.
Helder zijn over het gebruik en de gebruikers	Deelnemers aan de focusgroepen stellen vast dat ze dit in het begin moeten bepalen willen ze een goed gebruik van evaluaties kunnen garanderen. Dit is voor hen een zwak punt.
Beschikken over een planning van evaluaties	Voor actoren met M&E personeel gaat dit soms verder dan een gewone planning: het integreert ook de hierboven vermelde stappen. De andere organisaties hebben het meestal over een planning in de tijd.
Zorgen voor een link met het beleid van de organisatie	Deelnemers stellen dat een goede 'governance' van een evaluatie zorgt voor een gemeenschappelijke rode draad in de evaluaties en voor een link met het beleid en de beslissingen van de organisatie over strategie en programma's.
Zorgen voor een begeleiding van de evaluatie door een bredere groep (bijv. begeleidingscomité)	Er is geen echte consensus over wie allemaal in die groep moet zitten. De meeste organisaties zien dit vooral als een interne groep met deelnemers uit de organisatie zelf. Een minderheid van organisaties gaf aan ook externen te (willen) betrekken, in veel gevallen ging het dan om de dossierbeheerder van DGD.
Verzekeren van evalueerbaarheid en een duidelijke link met de monitoring	Over deze stap was minder consensus en duidelijkheid bij de verschillende deelnemers. Vooral organisaties met M&E personeel zagen het belang van deze stap.

Heldere referentietermen die verwijzen naar de criteria van de OESO (DAC comité)	Er was geen echte consensus over de inhoud van de referentietermen. Sommige organisaties hechtten bijzonder belang aan de precisering van procedures. Zeker voor BTC en organisaties met een koepelfunctie bleek dit van belang. De meerderheid van de organisaties legt vooral nadruk op 'technische' referentietermen die aangeven wat van de evaluator verwacht wordt.
--	---

In de beoordeling van hun eigen praktijk, stellen de actoren dat *de fase van het ontwerp/ontwikkeling* van de evaluatie redelijk is: de actoren vinden dat ze werk maken van duidelijke referentietermen, dat die beantwoorden aan de formele criteria van de OESO en verwijzen naar de DAC-criteria van relevantie, efficiëntie, doeltreffendheid, duurzaamheid en impact. De actoren geven ook aan dat de afbakening van de scope, doelstellingen en evaluatievragen veel beter kan.

Op basis van een kruising van deze informatie met de literatuur en informatie over de screening, stellen we volgende aandachtspunten vast over de **ontwerpfase van evaluaties**:

- Evaluatie en integratie in de programmacyclus: slechts een minderheid van organisaties denkt na over evaluaties bij de formulering van de programma's. Vaak wordt pas nagedacht over de aard en inhoud van de evaluatie bij de opstelling van de referentietermen terwijl het natuurlijk belangrijk is om hierover veel eerder in de programmacyclus reflectie te voeren, zodat bijv. ook verzekerd is dat het monitoringsysteem van bij de start van de uitvoering de juiste data verzamelt.
- Actoren maken eerder beperkt gebruik van deze fase om de legitimiteit van de evaluatie en het draagvlak ervan binnen de organisatie te versterken: er is weinig specifieke aandacht voor deze fase en de uitvoering gebeurt vaak onder tijdsdruk (die soms een gevolg is van slechte planning).
- Nadenken over het type evaluatie: het ontwerpen van evaluaties gebeurt weinig dynamisch en er is onvoldoende reflectie over welk type evaluatie wenselijk is en waarom. Er is weinig variatie in opzet. Dit wordt deels bevestigd door de DGD-screening: de zwakste scores (in het domein resultaatsgericht beheer) betreffen het uitvoeren van meta-evaluaties, transversale evaluaties over verschillende landen heen en peer-evaluaties (hoewel deze bijzonder interessant geacht worden door de actoren). Deze zwakkere scores worden bevestigd door de focusgroepen en de input van de ANGTS in deze focusgroepen: het is geen gangbare praktijk.
- Evalueerbaarheid¹⁴ en ToC: er is weinig praktijk tot op heden in het gebruiken van een ToC om te bepalen wat geëvalueerd kan worden en wat relevant is om te evalueren, organisaties beseffen wel dat de ToC kan helpen om te komen tot een betere formulering van outcomes en tot het formuleren van goede evaluatievragen. De DBE-studie rond evalueerbaarheid (2016) toont aan dat evaluatie op het niveau van outcomes en de link met impact eerder zwak is. Dezelfde studie geeft ook aan dat organisaties die effectief aan de slag gaan met een ToC vaak sterker staan op het vlak van evaluatie.
- Evalueerbaarheid en dataverzameling: gepaste datacollectie om evaluatievragen te kunnen beantwoorden of het verifiëren daarvan gebeurt niet systematisch bij de actoren, bijv. bij de start van het programma en het opstellen van de ToR of bij de start van de evaluatie. De link tussen indicatoren en gepaste methoden voor

¹⁴ De term 'evalueerbaarheid' verwijst naar de theoretische en praktische mogelijkheid om een interventie te evalueren op een betrouwbare en geloofwaardige manier. Is de onderliggende analyse van de gewenste verandering duidelijk, zijn indicatoren helder (met name op het niveau van de outcome), is er een baseline die gegevens verzamelt en is er een gepast systeem voor datacollectie, consolidatie en stockage, ... ?

dataverzameling, de verwerking van (kwalitatieve) data en het stockeren van gegevens zodat ze nadien makkelijk te exploiteren zijn, is voor veel actoren nog een verbeterpunt. De controle van de kwaliteit van dataverzameling op het terrein is niet altijd verzekerd en veranderingen in de interventie worden niet altijd gevolgd door aanpassingen in de monitoring en datacollectie (zie de DBE-studie rond evalueerbaarheid). Het verzamelen van de baselinegegevens bij de start van het programma is voor veel organisaties een heuse uitdaging en staat vaak onder tijdsdruk, maar het gebeurt wel met meer aandacht dan vroeger.

- Het identificeren van een realistische en relevante evaluatiescope en het formuleren van goede evaluatievragen (helder, niet te veel, relevant, ...) is een uitdaging voor de actoren. Dikwijls verwijzen de actoren naar de DAC criteria die goed gekend zijn: voor velen vormt dit het voornaamste kader voor een evaluatie. Recente gesprekken binnen de Wereldbank en de OESO DAC¹⁵ en de DBE-studie rond evalueerbaarheid (2016) geven echter aan dat niet elke evaluatie noodzakelijk op alle DAC-criteria moet ingaan en dat evaluaties ook andere evaluatievragen naar voor kunnen schuiven. De bovenvermelde DBE-studie benadrukt ook dat de afbakening en beschrijving van het object van evaluatie, meer bepaald de doelgroep (waar moeten/kunnen effecten gemeten worden?) over het algemeen nog een zwak punt is en dit werd bevestigd door de focusgroepen.
- Afstemmen van het beschikbare budget, de (doorloop)tijd die beschikbaar is en de scope/evaluatievragen (verwachtingen) gebeurt niet systematisch in de praktijk.
- De zwakste scores van de screening (in het domein resultaatsgericht beheer) betreffen onder meer de organisatie van **stuurgroepen** voor het beheer van de evaluaties, met daarin een vertegenwoordiging van partners en/of externe actoren: dit is doorgaans niet de praktijk bij een meerderheid van de actoren wat werd bevestigd door de focusgroepen.

Stappen die belangrijk zijn in een evaluatieproces – de mening van de ANGS	
Uitvoering van een evaluatie	
Selectieproces van de consultant	In veel organisaties gebeurt dit eerder informeel. In drie focusgroepen was er veel aandacht voor de juridische aspecten en het respecteren van de wet op overheidsopdrachten. De onafhankelijkheid van de evaluator zorgde steeds voor veel discussie onder de deelnemers: onafhankelijkheid is belangrijk maar actoren schuiven andere elementen naar voor die voor hen even belangrijk zijn: houding, ervaring, kennis, ...
Briefing van de consultant	In drie focusgroepen was er veel aandacht voor de coherentie tussen verwachtingen en benaderingen en voor een verfijning van de methodologie. Niet alle actoren vinden dat hiervoor een startrapport (of <i>inception report</i>) nodig is. In één focusgroep vond men dat deze briefing ook in moest gaan op de evalueerbaarheid (actoren met M&E personeel).
Praktische voorbereiding (logistiek)	Organisaties geven aan dat hun partnerorganisaties hier een belangrijke taak krijgen.
Uitvoering van de missie/evaluatie	Slechts enkele actoren zien voor zichzelf als opdrachtgever van de evaluatie een rol tijdens de uitvoering van de

¹⁵ Zie bijv. <https://ieg.worldbankgroup.org/news/conversations-future-development-evaluation>.

	evaluatie, zoals het verzekeren van een geregeld contact (met het oog op verzekeren van kwaliteit en leren). Organisaties hechten belang aan een participatieve aanpak en aan een evaluator die dit ondersteunt en een sfeer van vertrouwen en openheid kan creëren. Verschillende actoren onderlijnen dat zij graag en vaak met nationale consultants werken.
Restitutie op het terrein en op de zetel van de organisatie	Actoren vinden dit belangrijk om ervoor te zorgen dat er eigenaarschap (<i>ownership</i>) is over de resultaten en aanbevelingen van de evaluatie en om een realiteitscontrole te doen van de aanbevelingen. Voor de deelnemers in 2 focusgroepen begint hier al het gebruik van de evaluatie.
Commentaar op het rapport	Hier blijken zeer verschillende benaderingen voor te zijn. Slechts 1 focusgroep ging in op het aspect van betrouwbaarheid en benadrukte dat dit afhangt van de bronnen die gebruikt worden, de trackrecord van de evaluator, de triangulatie van informatie, het verzekeren van feedback van actoren op het terrein, de methode die werd gebruikt.
Controle van de kwaliteit van het rapport	De voornaamste criteria volgens de deelnemers zijn (niet noodzakelijk voor alle organisaties in dezelfde volgorde van belangrijkheid): beantwoordt het rapport aan de referentietermen, vertelt het rapport ons iets nieuw, is het concreet en synthetisch (niet te lang, aangename lectuur), is het constructief, zijn de aanbevelingen realistisch, duidelijk en relevant, is de informatie betrouwbaar?

In het algemeen, beoordelen de actoren de *uitvoering* van de evaluatie als redelijk sterk; ze baseren zich dan voornamelijk op de aspecten van logistieke organisatie en op de briefing van de consultants bij aanvang van de evaluatie.

Kruising van informatie leidt tot de volgende vaststellingen inzake aandachtspunten in de uitvoering van de evaluatie:

- Rol van de organisatie tijdens de uitvoering: de actoren zien voor zichzelf vaak geen echte rol tijdens de uitvoering van de evaluaties tenzij voor de briefing, de restitutie en het faciliteren van de logistieke aspecten (met de partners). Dit is opvallend omdat er tijdens de uitvoering natuurlijk veel mogelijkheden zijn om te leren, op te volgen en bij te sturen in functie van een goed gebruik van de evaluatie en de kwaliteit van het evaluatierapport. Actoren schuiven de uitvoering soms te makkelijk door naar de evaluatoren zonder veel stil te staan bij hun rol.
- Verzekeren dat de evaluator zijn methodologische benadering verfijnt of bevestigt vooraleer de evaluatie start: dit wordt niet altijd systematisch verzekerd en actoren argumenteren dat het bij eenvoudige evaluaties niet altijd noodzakelijk is. Een inception report of een nota waarin de definitieve evaluatieaanpak wordt beschreven en onderbouwd, wordt door een groot deel van de actoren niet opgelegd of gevraagd aan de evaluator.
- Het organiseren van een restitutie op het terrein blijkt voor veel actoren een gangbare praktijk; wat evenwel moeilijker is, is het verzekeren van de betrokkenheid van de partners bij de evaluatie na de missie (dat heeft ook een invloed op het gebruik van de evaluatie door de partners (zie hieronder)).
- Analyse van de methodische kwaliteit van een evaluatierapport: de meerderheid van de actoren doet dit niet op een systematische manier en heeft hier geen tools voor. De

meeste actoren kijken wel naar volgende aspecten (en analyseren dat op een niet-systematische manier): beantwoordt dit rapport aan de referentietermen, is er een goede analyse, leren we er iets uit, zijn de aanbevelingen haalbaar/relevant?

Stappen die belangrijk zijn in een evaluatieproces – de mening van de ANGS	
Gebruik van een evaluatie	
Discussie over de aanbevelingen	De meeste actoren organiseren geen specifieke discussie of workshop met de diverse stakeholders over aanbevelingen, maar nemen dit mee in hun commentaar op het rapport. In 1 focusgroep benadrukten deelnemers het belang om tijd te nemen om dit door te praten met de partners.
Verspreiding van het rapport	De organisaties geven aan dat het belangrijk is om het rapport intern en extern te verspreiden. Een beperkt aantal organisaties vindt het ook belangrijk om de eindbegunstigden daarbij te betrekken.
Management response	Dit werd in drie focusgroepen aangehaald als een belangrijke stap, hoewel er discussie was over de graad van formaliteit, de inhoud en het niveau waarop dit geformuleerd wordt.
Publicatie van het rapport en de management response	De actoren willen een onderscheid maken tussen verspreiding van het rapport (intern, extern, stakeholders) en de publicatie ervan.
Uitvoering van de aanbevelingen/operationalisering	Alle organisaties geven aan dat dit moet gebeuren. Er is echter geen consensus over de manier waarop: er zijn verschillende ideeën, gaande van aantekeningen tijdens een vergadering tot een volwaardig actieplan waarvan de uitvoering en de resultaten ook systematisch opgevolgd worden.

De fase van *exploitatie/gebruik van de evaluatie* wordt door de actoren zelf gezien als de zwakste fase in het proces (vaak informeel, impliciet, globaal minder aandacht, ...) en vooral het concreet werken met de aanbevelingen is weinig ontwikkeld bij de meeste actoren. Ze geven zelf aan dat de exploitatie zwak is omdat het gebruik van resultaten (hoe, voor wie, op welke manier, ...) bij de start van de programmacyclus en het concipiëren van evaluaties te weinig geëxpliciteerd wordt. Opnieuw is duidelijk dat evaluatieprocessen dikwijls losstaan van de organisatie en haar strategische en programmacycli.

Op basis van de informatie uit de focusgroepen, de resultaten van de screening en de literatuur, komen we tot een overzicht van de voornaamste knelpunten/aandachtspunten in de fase van het gebruik van de evaluatie:

- De uitvoering van de aanbevelingen: alle organisaties geven aan dat ze goed gebruik willen maken van evaluaties en dat dit belangrijk is. De meerderheid van de actoren geeft aan dat ze nog niet aan de slag zijn gegaan met de aanbevelingen om diverse redenen: de scope van de evaluatie bleek achteraf niet de juiste, de aanbevelingen zijn niet relevant, of ze kunnen niet meer gebruikt worden omdat de nieuwe interventies al geformuleerd zijn, er is geen tijd om veel aandacht te besteden aan de aanbevelingen, er zijn weinig incentives in de organisatie om het te doen (bijv. verantwoordelijken voor de evaluatie worden er niet op aangesproken), er was tijdens

de evaluatie te weinig aandacht voor de noden van de organisatie en voor het leren, ... De meeste actoren kunnen geen 'actieplan' voorleggen waarin duidelijk beschreven staat hoe rekening zal gehouden worden met de aanbevelingen. De actoren nuanceren evenwel: een actieplan (of management response) is niet altijd nodig om aanbevelingen op te nemen, aanbevelingen kunnen ook op een andere manier doorwerken en dit wordt vaak mogelijk gemaakt door de specifieke aandacht van individuele personen binnen de organisaties voor aanbevelingen.

- Gebruik voor de voorbereiding van nieuwe interventies: een aantal actoren maakt systematisch gebruik van evaluaties om nieuwe programma's voor te bereiden, maar door de vaak laattijdige uitvoering van evaluaties is dat bij de meerderheid van de organisaties een echte uitdaging. Dit wijst er opnieuw op dat evaluaties weinig systematisch geïntegreerd zijn in de programmacyclus.
- Weinig actoren hebben tot op heden de praktijk om een formele management response te schrijven.
- Publicatie van het rapport en de management response op de website is geen gangbare praktijk en stoot op heel wat vragen bij de actoren. De meerderheid van de actoren is van mening dat het niet altijd opportuun is om het volledige rapport op de website te publiceren (bijv. omwille van veiligheidsredenen). De meeste actoren zijn bereid om te onderzoeken wat de gepaste formats zijn om evaluaties en evaluatieresultaten te delen met een breder publiek.
- Uit de discussies in de focusgroepen kwam niet automatisch naar voor dat organisaties de conclusies van hun evaluaties willen delen met andere actoren (of op welke manier ze dit doen), dit was voor hen geen prioriteit. Slechts enkele actoren opperden tijdens 2 focusgroep discussies dat evaluaties gebruikt kunnen worden om de samenwerking/het leren binnen de GSK's (Gemeenschappelijke Strategische Kaders) te voeden.

De analyse van de evaluatiepraktijk en -processen leidt tot een aantal conclusies die raken aan de capaciteit van organisaties om de evaluatiefunctie te organiseren:

- Betrokkenheid van de partners in de verschillende fasen van de evaluatie: dit kwam niet automatisch naar voor als een belangrijk punt in de focusgroepen. Zoals gevraagd, blijkt dat de rol van de partners heel verschillend kan zijn afhankelijk van de scope van de evaluatie (worden ze zelf geëvalueerd of evalueren ze de actor?) en dat die rol vaak niet expliciet is gedefinieerd. Hun rol kan het volgende omvatten: aanleveren van data, opstellen van referentietermen, optreden als respondent/interviewee tijdens de terreinmissie, reacties geven op het rapport en de aanbevelingen, ... Evaluaties gebruiken om de capaciteit van partners rond M&E te versterken is geen gangbare praktijk en/of gebeurt impliciet. De studie rond evalueerbaarheid gaf als knelpunt aan dat de M&E systemen van de actoren weinig gearticuleerd/geïntegreerd waren met de systemen van de partners: gezien de zwakke capaciteit van de partners om een goede datacollectie te verzekeren is dit een belangrijk aandachtspunt.
- Zich laten omringen door externe experts voor de begeleiding van de evaluatie (bijv. door middel van stuurgroepen en begeleidingscomités) om kwaliteit, transparantie en bruikbaarheid van evaluaties te verzekeren, gebeurt eerder beperkt.
- De governance van de evaluatie: het strategisch inzetten van evaluaties en het linken van evaluaties met andere processen binnen de organisatie (waken over coherentie), het ontwikkelen van een interne evaluatiedynamiek, het betrekken van Raad van Bestuur, ... zijn minder sterk ontwikkeld.
- De technische competenties bij het personeel inzake evaluatie: minimale kennis over evaluatie om ervoor te zorgen dat de scope van de evaluatie goed wordt afgebakend,

dat de juiste evaluatievragen worden gesteld, dat de methoden die gebruikt worden door de evaluator goed begrepen worden, dat de evaluator kan bevraagd worden over zijn methoden, dat de kwaliteit van een rapport ten aanzien van betrouwbaarheid en validiteit systematisch geanalyseerd wordt, ... is over het algemeen weinig verspreid binnen de organisaties.

- De (meerjaren-) planning van evaluatie(s) en budgettering: de planning van alle evaluaties bij de start van een programma (het DGD-programma in het bijzonder) waarbij grondig wordt nagedacht over het nut van de evaluatie, het type, de gepaste methodologie, ... is geen gangbare praktijk voor de meeste actoren: er bestaan beleidsnota's die de grote oriëntaties op het vlak van evaluatie vastleggen maar die zijn vaak onvoldoende operationeel. Dit vraagt uiteraard tijd: veel actoren moeten daar nu mee starten, hoewel het bij de start van de uitvoering van een programma vaak moeilijk blijkt om een focus op M&E te houden en erin te investeren. Veel ANG's budgetteren 1% van het programmabudget voor evaluaties omdat dit verplicht is: het is veelal niet duidelijk in welke mate dit budget de reële kost weerspiegelt en wat de verhouding is tot het budget voor monitoring en datacollectie. Het budget voor monitoring is weinig 'leesbaar' in de budgetten. Uit de DBE-studie rond evalueerbaarheid kan worden afgeleid dat middelen vaak vooral geïnvesteerd worden in monitoring op output niveau en dat dit soms ten nadele is van middelen voor evaluatie.
- Relaties met evaluatiebureaus/consultants: organisaties geven aan dat het vaak moeilijk is om goede evaluatoren te vinden die een goede kennis hebben van het evaluatieonderwerp en beschikken over de gevraagde competenties. Vaak gaat het om een combinatie van expertise in evaluatie, het faciliteren van leren en thematische kennis. Het consortium stelt vast dat de meeste organisaties slechts contact hebben met een kleine pool van consultants en niet proactief zijn om deze pool te vergroten. Soms zijn de verwachtingen ten aanzien van het profiel van de evaluator weinig realistisch. Voor de actoren is de onafhankelijkheid van de evaluator belangrijk maar zeker niet het belangrijkste: zijn capaciteit voor analyse en kritische/constructieve ingesteldheid zijn dat evenzeer. De meta-evaluatie van 2013 in opdracht van de DBE geeft echter aan dat er een duidelijke correlatie is tussen onafhankelijkheid en kwaliteit van de rapporten en dat dit dus toch sterker op de voorgrond moet staan.
- Samenwerking met andere organisaties: het uitwisselen van informatie, ondersteunen van evaluaties (bijv. in stuurgroepen), gezamenlijk evalueren, ..., is weinig ontwikkeld.

De knelpunten/aandachtspunten die hierboven worden beschreven, kunnen in bepaalde gevallen leiden tot niet-realistische, weinig relevante of minder bruikbare referentietermen en evaluaties. Het belang van deze knelpunten wordt bevestigd door studies. In bepaalde gevallen geven die studies ook aan wat dan wel beter zou werken. Een studie van ITAD (Information Technology and Agriculture Development) onderlijnt bijv. het volgende: *"The MEL (monitoring, evaluation, learning) systems that are most useful, according to the NGOs using them, combine long-term relationships with partner NGOs with the ability to do data analysis close to the ground and MEL capacity building of local partners. They also have deep integration of MEL within an NGO head office, ensuring it is perceived as a collective responsibility, and a focus on improving work with beneficiaries and partners, rather than proving effectiveness to donors or external stakeholders"*.¹⁶

Eén organisatie vat het risico samen als volgt tijdens een focusgroep discussie: *'on risque de suivre seulement les exigences du bailleur et son focus et non nos besoins en évaluation'*. Dit risico is niet onvermijdelijk: de huidige verwachtingen van DGD ten aanzien van evaluatie bepalen voor een groot deel, zeker voor kleinere organisaties (met een beperkt budget), de

¹⁶ ITAD (2014) *Investing in monitoring, evaluation and learning - Issues for NGO's to consider*.

evaluaties die zij zullen/kunnen uitvoeren en de momenten waarop die moeten gebeuren. De actoren kunnen echter zelf proactief aansturen op bruikbare evaluaties van goede kwaliteit en niet alleen omdat er een (terechte) vraag is naar verantwoording vanuit de overheid. Als ze daarbij rekening houden met bovenstaande aandachtspunten kunnen ze al een heel stuk verder komen.

Over BIO en BTC als specifieke actoren : uit de interviews en de documenten van BIO en BTC die geraadpleegd werden tijdens de eerste fase van dit onderzoek blijkt het volgende:

BIO (de Belgische Investeringsmaatschappij voor Ontwikkelingslanden) neemt participaties in of verstrekt lange termijn financiering aan private bedrijven, financiële instellingen, fondsen en investeringsprojecten. Daarnaast verleent ze technische assistentie om de financiële leefbaarheid en ontwikkelingsrelevantie van investeringsprojecten te versterken. BIO heeft de voorbije 3 jaren belangrijke inspanningen geleverd om de informatie en opvolging van haar investeringsprojecten te versterken en om de impact van haar investeringen op ontwikkeling op te volgen. De organisatie voert jaarlijks evaluaties uit van 5 cases. Die evaluaties lieten toe om lessen te trekken uit een specifieke sector of regio en het beleid ervan aan te passen. De inspanningen van de organisatie zijn behoorlijk sterk vergeleken met die van de meerderheid van de andere Development Finance Institutions (DFI's): minder sterk dan die van de grote maar beduidend meer dan kleinere en middelgrote DFI's. Een specifieke uitdaging is een meer regelmatige, systematische analyse van de beschikbare monitoringgegevens ten aanzien van ontwikkelingseffecten: dit zou toelaten om trends en verbanden te identificeren en sneller lessen te trekken uit zowel de huidige als de toekomstige investeringen. Deze analyse van de informatie uit het monitoringsysteem kan onder meer helpen om de juiste evaluatievragen op te stellen en de conclusies van de casestudies te vergelijken met de algemene vaststellingen per sector, land... Gezien de regelmatige financiële samenwerking tussen de verschillende DFI's en het overleg/de uitwisselingen op Europees niveau, lijkt het aangewezen om verder te onderzoeken hoe haar 'peers' betrokken kunnen worden in het certificeringssysteem van BIO. Het lijkt bovendien zinvol dat de Belgische overheidsactoren het thema van de evaluatie van de DFI's op de agenda zetten tijdens hun overleg/uitwisselingen met hun Europese collega's.

BTC (de Belgisch Technische Coöperatie) wordt omgevormd tot ENABEL, het ontwikkelingsagentschap van de Belgische overheid dat in een beheerscontract met de overheid vanaf januari 2018 zal werken met een nieuw mandaat en een daaraan aangepaste governance structuur. Meer autonomie, meer verantwoordelijkheid en een grotere rol bij de start van de projectcyclus staan hierbij centraal. Deze hervorming geeft BTC/ENABEL meer flexibiliteit. De combinatie van meer autonomie en flexibiliteit gaat gepaard met hogere eisen ten aanzien van kwaliteit en dat heeft een invloed op monitoring en evaluatie: bijv. belang om de onafhankelijke interne evaluatiedienst verder uit te bouwen en naast twee standaard externe evaluaties op projectniveau (de tussentijdse en de eindreviews), de M&E dusdanig te organiseren dat er in de toekomst meer aandacht is voor strategische evaluaties en evaluaties van de kwaliteit van de instrumenten. Anderzijds is het belangrijk dat de feedbackloops over resultaten korter worden en dat lessen sneller getrokken worden (met andere woorden aandachtig zijn voor interne evaluatie die korter op de bal moet spelen). BTC heeft de laatste jaren fel geïnvesteerd in het uitwerken van een normatief en methodologisch kader voor evaluaties en in de steun aan BTC-medewerkers om te evolueren naar een meer coherente aanpak. Dat gebeurde in eerste instantie in het kader van de projectcyclus en nadien ook doorheen de hele organisatie en op een meer strategisch niveau. Die inzet beantwoordt aan twee bekommernissen die voor de overheid belangrijk zijn: (i) meer focus op outcomes en 'the missing middle' tussen outcome en impact en (ii) investeren in betrouwbare evaluatierapporten. Uit het evaluatiecharter en het daaruit voortvloeiende actieplan voor 2017 blijkt dat BTC de ambitie heeft om daar verder in te investeren.

2.1.2. Capaciteit van de andere stakeholders

Om te kunnen bepalen hoe het instrument in de praktijk kan werken, is het belangrijk om te kijken naar de capaciteit van de diverse stakeholders, maar ook naar de rol die ze spelen in de Belgische ontwikkelingssamenwerking. We hebben het dan met name over de actoren zelf, DGD (en meer bepaald D1, D2, D3 en D4¹⁷), de DBE en de Strategische Cel voor Ontwikkelingssamenwerking.

Op basis van interviews en documentstudie kunnen deze rollen als volgt samengevat worden:

DGD:

- moet informatie verwerken (de federale overheidsdienst is aan het evolueren naar een kenniscentrum);
- onderzoekt de capaciteit van de actoren op vlak van resultaatsgericht beheer en spreekt hen op hun verantwoordelijkheid aan;
- neemt beslissingen over de financiering van nieuwe programma's/interventies op basis van de conclusies van evaluaties;
- draagt bij tot de kwaliteit van de referentietermen van de ANGS, bijv. door het formuleren van aandachtspunten of evaluatievragen;
- D1 en D2 kunnen samenwerking (onder meer rond evaluatie) op Europees niveau initiëren/stimuleren, bijv. voor BTC en BIO.

De DBE en D4 hebben elk een rol in het ondersteunen van kwaliteit en het bijdragen tot harmonisatie in de evaluatie aanpakken en in het ondersteunen van het gebruik van evaluaties.

De ANGS, hun federaties, BIO en BTC moeten de evaluatiefunctie garanderen alsook de kwaliteit van de evaluatieprocessen. Zij moeten sterk aansturen op de kwaliteit van externe evaluatoren en hun rapportage¹⁸ en waken over het gebruik van de evaluatieresultaten.

De federaties ondersteunen hun leden en bieden vorming en capaciteitsversterking op het vlak van evaluatie.

Op het vlak van de capaciteit van de verschillende stakeholders, moeten volgende zaken bemerkt en herinnerd worden:

- DGD, meer bepaald D3, D1 en D2: de technische kennis op het vlak van evaluatie is beperkt, alsook de capaciteit om een oordeel te vellen over de uitvoering en het resultaat van evaluaties. De capaciteit (vooral in termen van tijd en human resources) om evaluatierapporten op een systematische manier te verwerken en ze te gebruiken in functie van besluitvorming is eveneens beperkt.
- DGD, meer bepaald D4: er is veel ervaring omtrent evaluatie in deze dienst en toegang tot allerhande kennis. De capaciteit in termen van human resources is echter zeer beperkt alsook de mogelijkheid om processen van capaciteitsversterking binnen het ministerie (meer bepaald D3) te beïnvloeden en te ondersteunen.
- De DBE heeft ervaring omtrent evaluatie in huis maar het aantal beschikbare HR is eerder beperkt. Er zijn financiële middelen voor de uitvoering van het instrument in een eerste cyclus maar die houden voorsnog weinig rekening met de noodzaak om

¹⁷ D1 is de Geografische Directie, D2 de Thematische Directie, D3 de Directie Civiele Maatschappij en D4 de Directie Organisatiebeheer

¹⁸ De actoren moeten aansturen op kwaliteit van de externe evaluatoren en hun rapporten, het is aan de evaluatoren om de kwaliteit van het rapport te verzekeren.

ook mensen binnen de dienst beschikbaar te stellen voor de uitvoering en het beheer van het instrument. Daarenboven bemoeilijkt de onafhankelijkheid van de dienst het verlenen van concrete ondersteuning bijv. aan de andere departementen binnen de federale overheidsdienst.

- Federaties: zij willen investeren in hun eigen kennis en wensen ondersteuning aan te bieden aan de actoren. De federaties hebben intussen met hun leden reeds werkgroepen opgericht rond M&E. De operationalisering hiervan moet nog verduidelijkt worden (alook het budget dat hiervoor ter beschikking is).
- De actoren (de ANGS, BTC en BIO): er zijn meerdere actoren die specifieke en interessante benaderingen hebben ontwikkeld omtrent evaluatie en die dus in staat zijn om op te treden als *peers* en om informatie en kennis te delen. Er is bij de actoren echter weinig ruimte om effectief extra (financiële en menselijke) middelen te investeren in verbetertrajecten die zich enkel richten op evaluatie; er zijn zoveel andere taken die vervuld moeten worden. Geen enkele actor is ertegen om samen met andere actoren te leren, maar de ruimte daarvoor lijkt wel beperkt en hierin investeren is voor de meeste actoren geen prioriteit.
- In het geheel van de sector is er op dit moment nog steeds te weinig gedeeld begrip over wat een evaluatie precies is en wat er allemaal belangrijk is. Dit bemoeilijkt de (inhoudelijke) dialoog tussen actoren en stakeholders.

2.1.3. Informatienoden van de actoren en stakeholders

Als het over evaluaties en evaluatierapporten gaat, waar hebben de verschillende stakeholders dan nood aan? Waar moet het instrument toe bijdragen? Hieronder vatten we samen wat blijkt uit de interviews en de focusgroep gesprekken. Hieruit kunnen we concluderen, hoewel er soms verschillende accenten worden gelegd, dat er toch overeenkomsten zijn tussen de uitgedrukte noden, onder andere wat de nood om te kunnen beschikken over betrouwbare informatie over resultaten en veranderingen betreft:

De ANGS, BTC, BIO en de federaties willen evaluatierapporten die:

- beantwoorden aan de referentietermen;
- constructief zijn in hun analyse (en rekening houden met de context van de interventies);
- helder en beknopt zijn;
- betrouwbare informatie bevatten;
- een kritische analyse bevatten;
- iets nieuws vertellen;
- realistische, heldere en relevante aanbevelingen hebben.

De ANGS hebben in het proces nog bijkomende noden uitgedrukt : evaluatierapporten moeten:

- informatie geven over resultaten en verandering (en de verschillen daarin bij de verscheidene doelgroepen);

- informatie geven over de bijdrage van de organisatie en van haar strategie aan de resultaten;
- verduidelijken wat de link is tussen outcome en het impact niveau;
- verduidelijken *hoe* resultaten werden behaald;
- toelaten te leren over thema's (inhoud) en gevoerde strategieën (aanpak).

De stakeholders van de administratie onderstrepen vooral het belang van betrouwbare rapporten die:

- getoetst werden aan de perceptie van hun mensen op het terrein (D1);
- aangeven of de resultaten werden bereikt (en hun bijdrage aan de realisatie van de sustainable development goals (SDG's) (D1));
- informatie geven over impact (relatie tussen outcome en impact (D4); effecten op ontwikkeling (D2));
- verduidelijken wat werkt en wat niet;
- een duidelijke link hebben met de strategische beslissingen van de ANGS, BTC en BIO.

De DBE heeft nood aan betrouwbare rapporten voor de meta-evaluaties en wil een beter zicht krijgen op hoe monitoring wordt georganiseerd in functie van evaluatie.

Tenslotte willen de federaties zicht krijgen op de knelpunten, goede praktijken, methodes en instrumenten die gebruikt worden door de actoren.

2.2. Bevindingen vanuit de literatuur

De beknopte literatuurstudie (zie overzicht van geconsulteerde documenten in bijlage, met de weblinks) maakt snel duidelijk dat een instrument voor de certificering van evaluatiesystemen, zoals gevraagd in de referentietermen, geen voorbeeld kent in andere landen of sectoren. Er bestaan wel verschillende instrumenten in de sector van ontwikkelingssamenwerking die aandacht hebben voor de kwaliteit van evaluaties. Het gaat dan bijv. over instrumenten die toelaten om de kwaliteit van rapporten systematisch te evalueren¹⁹, of die de capaciteit willen versterken van die mensen die evaluaties moeten aansturen en beheren.²⁰²¹

¹⁹ Zie bijv. de checklist voor evaluatie van methodische kwaliteit van het IOB (Dienst Inspectie en Ontwikkelingssamenwerking) in Nederland. Of die van de ILO (2014) of de 'quality assessment system for decentralized evaluation reports' van de UNDP.

²⁰ Zie UNEG (2015): review van een competency framework en noodzakelijke professionele competenties voor evaluatie. DFID heeft in 2014 een analyse gemaakt over het inbedden van evaluatie in het Britse Ministerie voor internationale samenwerking. Die inbedding steunde op: verzekeren van voldoende staff met competenties inzake evaluatie, bepalen van standaarden voor 'evidence', focus op impactevaluatie en ontwikkeling van een professioneel kader van geaccrediteerde evaluatiespecialisten.

²¹ Het OESO rapport (2016) *Evaluation systems in Development Cooperation*, geeft een overzicht van wat verschillende landen doen om de kwaliteit van hun evaluaties te verzekeren. De volgende aspecten worden belicht: evaluatiesysteem en governance, evaluatieprocessen en doeltreffendheid van evaluaties.

Er wordt in de literatuur frequent verwezen naar **de voorwaarden die vervuld moeten worden om de kwaliteit van evaluaties te garanderen**, zoals: het belang van leiderschap binnen een organisatie; het beschikken over voldoende middelen en mensen (wat vaak weinig expliciet in de budgetten naar voor komt) en over mensen met de juiste competenties (op het terrein en de zetel).²² Ook verwijst de literatuur naar de stimulerende rol die een subsidiërende overheid kan spelen bij het promoten van inspanningen om de evaluatiepraktijk van organisaties te versterken (Carter 2013²³; Mackay 2009²⁴).

Het is voorts niet moeilijk om in de literatuur normen terug te vinden die de kwaliteit van evaluaties of M&E bepalen.²⁵ Maar tegelijkertijd wordt beklemtoond hoe problematisch het is om in de praktijk aan veel van die normen te voldoen. De organisatie en aansturing van de evaluatiefunctie binnen een organisatie worden namelijk sterk bepaald door haar context, mandaat en visie op verandering. Normen die stellen dat de onafhankelijkheid van de evaluatiefunctie enkel gegarandeerd kan worden door een evaluatiedienst die losstaat van de operationele departementen van de organisatie houden geen rekening met de realiteit van bijv. kleinere organisaties die de onafhankelijkheid noodgedwongen op een andere manier (moeten) organiseren (bijv. door een referentiegroep met externe experts).

Er bestaat veel ondersteunend materiaal voor organisaties die willen investeren in de kwaliteit van hun evaluaties. De lezer kan hiervoor terecht in de bijlagen (overzicht van geconsulteerde documenten).²⁶

Interessant vanuit **een leegericht perspectief** zijn de studies van Lennie (2015) en Preskill & Boyle (2008) die aandacht besteden aan hoe organisaties hun evaluatiecapaciteit binnen de organisatie kunnen versterken. De studie van EBA (Expert Group for Aid Studies (Swedish external evaluation office))(2017), *Confronting the contradiction*²⁷, biedt interessante tools om beter om te gaan met het spanningsveld tussen leren en verantwoording afleggen dat inherent is aan veel evaluaties en niet altijd toelaat om een goede focus te bepalen voor de externe evaluaties.²⁸ Een benadering die focust op het belang van leren en continue verbetering sluit goed aan bij de huidige visie en trends rond monitoring- en evaluatiesystemen (Stern et al. (2012)²⁹, Van Ongevalle et al. (2014)³⁰, F3E (2017)³¹).

²² Zie ITAD (2014) *Investing in monitoring, evaluation and learning. Issues for NGO's to consider*. Deze studie is gebaseerd op wat NGO's zelf aangeven als belangrijk. Hieruit kwamen drie punten naar voor: (i) strategic investment and funding, (ii) leadership en (iii) adequate staff capacity. Ook interessant is Hallam, A & Bonino, F (2013) met inzichten uit het netwerk van organisaties die in de humanitaire sector actief zijn (ALNAP).

²³ Carter, B. (2013) *Evaluation Capacity Development. GSDRC, Applied Knowledge Services*. <http://www.gsdr.org/docs/open/hdq996.pdf>

²⁴ Mackay, K. (2012) *Conceptual framework for monitoring and evaluation*. In Lopez - Acevedo, G., Krause, P. and Mackay, K. (Eds.) (2012). *Building better policies. The nuts and bolts of monitoring and evaluation systems*. Washington, D.C.: World Bank. <https://openknowledge.worldbank.org/handle/10986/6015>

²⁵ Op internationaal niveau kunnen we bijv. verwijzen naar documenten van de OESO (2010), UNEG (2015,2016).

²⁶ We verwijzen hier alvast nogmaals naar de DBE studie over evalueerbaarheid (2016) die veel verduidelijkt over het belang om goede keuzes te maken inzake evaluatie (waarom evalueren, bepalen van scope, inbedding van evaluatie in de organisatie, ...). De titel 'geloofwaardig en zinvol evalueren', spreekt voor zich. Verder verwijzen we naar de website: www.betterevaluation.org die een zeer bruikbaar startpunt betekent voor iedereen die vragen heeft over evaluaties, en naar de Rainbow Framework blog post van Irene Guijt die een goed overzicht biedt van hoe je best een evaluatie organiseert. Het framework kan gebruikt worden om evaluaties te plannen en te beheren maar ook om de kwaliteit van evaluaties te verifiëren en om de capaciteit van de organisatie te versterken op het vlak van evaluaties.

²⁷ <http://eba.se/en/confronting-the-contradiction-an-exploration-into-the-dual-purpose-of-accountability-and-learning-in-aid-evaluation/#sthash.iDgZrdDN.UICfh0ZU.dpbs>

²⁸ F3E in Frankrijk heeft veel gewerkt rond evaluatie en publiceerde in 2017 een analyse van de praktijken voor een transformatieve evaluatie onder de titel: *Comment renforcer les apprentissages et les changements à travers l'évaluation?*

²⁹ Stern E, Stame N, Mayne J, Forss K, Davies R and Befani B (2012) *Broadening the range of designs and methods for impact evaluation*. Working Paper no 38 for The Department for International Development, April. London, UK. URL (consulted 17 September 2013): <http://mande.co.uk/blog/>; wp-content/uploads/2012/06/2012-Stern-Working-Paper-38-final-April-20121.pdf

Binnen de context van ontwikkelingssamenwerking is vooral de Fair Trade sector vertrouwd met het concept van certificering. Het proces van certificering gebeurt, net als dit het geval is voor andere systemen van certificering, op vrijwillige basis, bijv. rond kwaliteit (bijv. ISO), milieu, zorg etc. en organisaties die beslissen hierin te investeren hebben meestal ook een duidelijke en concrete return (bijv. toegang tot de markt, imagoverbetering, ...). De principes van hun richtlijnen rond het ontwerpen van standaarden, het versterken van geloofwaardigheid en de evaluatie van de effecten van certificeringssystemen (leidt het tot betere praktijk?) waren nuttig voor het denkwerk rond een instrument gelinkt aan evaluatiebeleid en -praktijk maar bleken in de praktijk minder bruikbaar.³²

Andere systemen of labels waarmee de sector van ontwikkelingssamenwerking vertrouwd is, zijn die van het EFQM-model (European Foundation for Quality Management): de meerderheid van de ANGS hebben ervaring met EFQM als referentie voor kwaliteitsverbetering en worden hierin gesteund door de federaties. Het model van EFQM kijkt naar wat een organisatie doet en welke resultaten een organisatie behaalt. Het model wordt gebruikt om de realiteit van de organisatie af te toetsen aan een ideaal. Het model werkt als een spiegel die lacunes blootlegt en aantoont waarop een organisatie kan focussen om in de toekomst beter te beantwoorden aan de behoeften en verwachtingen van belanghebbenden.

In de praktijk bestaat het model uit een vragenlijst waarbij de organisatie haar sterke punten en verbeterpunten opsomt (= EFQM-zelfevaluatie). Het EFQM-model biedt een structuur om de juiste vragen te stellen. Tegelijk biedt het ook het kader om deze oefening regelmatig te herhalen. Het doel van deze oefening is om te komen tot een (korte!) lijst van verbeteringsprojecten en zo systematisch te werken aan een cultuur van continue kwaliteitsverbetering binnen de organisatie volgens een PDCA-cyclus (Plan, Do, Check, Adjust/Act). Het EFQM-model bleek een meer bruikbare en belangrijke informatiebron voor het denkwerk rond het instrument.

De meeste gekende systemen voor certificering van kwaliteit, alsook EFQM zijn holistisch en bekijken de organisatie als een geheel. Dat is een belangrijk verschilpunt met de focus van het instrument op externe evaluatie en monitoring.

³⁰ Van Ongevalle, J, Huyse, H., Van Petegem, P. (2014). *Dealing with complexity through "actor-focused" Planning, Monitoring & Evaluation (PME)*, Evaluation, The International Journal of Theory, Research and Practice, Vol. 20(4) pp. 447– 466. <http://evi.sagepub.com/content/20/4/447>

³¹ F3E (2017) *Evaluer les évaluations*

³² Bronnen van inspiratie zijn onder meer de publicaties van ISO, "de" autoriteit op het vlak van certificering en standaards: "How to write standards" <https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/how-to-write-standards.pdf>; "A Guide to Good Practice"; https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/casco_guide.pdf ; "Conformity assessment for standards writers - Do's and don'ts" <https://www.iso.org/publication/PUB100303.html>; "ISO 26000 and SDGs" https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/iso_26000_and_sdgs.pdf; Eu-Ecolabel zie <http://ec.europa.eu/environment/ecolabel/how-to-apply-for-eu-ecolabel.html> ; de koepel ISEAL Alliance (www.isealliance.org) waarvan onder meer FairTrade, Utz, Rainforest Alliance lid zijn.

2.3. Conclusies over de opzet en uitwerking van het instrument

Rekening houdend met de aandachtspunten die werden geïdentificeerd, de rollen, noden en capaciteit van de diverse stakeholders, de beschikbare literatuur en de context heeft het consortium keuzes gemaakt en voorgelegd aan de sector (middels sense-making workshops en de verschillende bijeenkomsten van het begeleidingscomité). Dit was geen lineair proces, documenten werden vaak herwerkt en verder gepreciseerd om rekening te houden met evoluerende inzichten:

Gezien :

- het instrument een meerwaarde moet hebben voor de actoren;
- leren en voortdurend verbeteren belangrijk zijn voor alle stakeholders en dit een proces is dat tijd vergt;
- de vraag naar een instrument waar zelfevaluatie terecht (gezien het belang van leren) een belangrijke rol speelt;
- het belang om te kunnen beschikken over kwaliteitsvolle evaluatierapporten (zowel voor de DBE, DGD als de actoren zelf) die niet enkel bijdragen tot het leren maar ook tot het afleggen van verantwoording over de interventies;
- de literatuur en bestaande studies aangeven welke elementen belangrijk zijn om te komen tot evaluaties van goede kwaliteit (en die dus richtinggevend zijn);
- de aandachtspunten die werden geïdentificeerd met de actoren tijdens focusgroepen en die aangeven waar er verbetering mogelijk is bij de actoren;
- de vraag naar een realistisch instrument (op vlak van timing en budget) dat rekening houdt met de diversiteit van de actoren;
- de beperkte ruimte bij een groot aantal actoren (om budget en mensen in te zetten voor evaluatie en te leren met andere actoren), maar ook bij D3 om evaluaties doeltreffend te gebruiken;

Zijn de volgende keuzes volgens het consortium relevant:

- Het instrument is een initiatief voor 'Evaluation Capacity Building': 'Evaluation capacity development involves the design and implementation of teaching and learning strategies to help individuals, groups, and organizations, learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of ECB is sustainable evaluation practice—where members continuously ask questions that matter, collect, analyze, and interpret data, and use evaluation findings for decision-making and action. For evaluation practice to be sustained, participants must be provided with leadership support, incentives, resources, and opportunities to transfer their learning about evaluation to their everyday work. Sustainable evaluation practice also requires the development of systems, processes, policies, and plans that help embed evaluation work into the way the organization accomplishes its mission and strategic goals.'³³

³³ Preskill & Boyle (2008), *A multidisciplinary model of evaluation capacity building*; Lennie (2015).

- Het accent ligt op de analyse door de organisatie van haar eigen praktijk (hoe doen wij dat?). Het is ook de organisatie die zelf haar verbeterplan vastlegt en eigenaar is van de analyse en het verbeterproces.
- De zelfevaluatie is een collectieve oefening: een brede participatie en discussie binnen de organisatie draagt bij tot het leren als organisatie en levert de nodige energie voor verandering.
- Het instrument moet inhoudelijk voldoende ondersteunend zijn: het instrument moet door zijn opzet en de tools die worden aangeboden de organisatie toelaten haar evaluatiepraktijk te structureren, kritische vragen te stellen en het pad naar verbetering aan te geven. Dit betekent dat het instrument voor zelfevaluatie voldoende moet uitleggen waar het over gaat en de analyse van de organisatie moet ondersteunen, bijv. door hulp te vragen.
- Elke organisatie die beantwoordt aan een minimumcapaciteit kan een vraag voor certificatie indienen voor de vastgelegde einddatum.
- De zelfevaluatie gebeurt aan de hand van een lijst van criteria die zijn opgedeeld in drie delen: criteria rond capaciteit, criteria rond het externe evaluatieproces (ontwerp, uitvoering en gebruik) en criteria rond de methodologische kwaliteit van evaluatierapporten. De criteria zijn multidimensionaal en omvatten verschillende elementen. Ze zijn niet normatief opgesteld; bijv. de criteria schrijven niet voor hoe rollen en taken omtrent evaluatie moeten verdeeld worden binnen een organisatie, maar stellen wel dat dit helder moet zijn en moet bijdragen tot onafhankelijkheid, een goede coördinatie van de evaluaties en tot hun kwaliteit. De criteria werden getoetst aan de literatuur.
- Gezien het weinig normatieve karakter van de criteria kan het instrument door een grote diversiteit aan actoren gebruikt worden.
- De lijst van criteria bevat een aantal criteria die deel uitmaken van de minimumcapaciteit van de organisatie inzake evaluatie. Organisaties moeten minstens deze basiscapaciteit hebben alvorens ze in aanmerking komen voor certificering. De evolutie naar dat basisniveau is de verantwoordelijkheid van de organisatie (die daarvoor mogelijk beroep moet doen op ondersteuning). Organisaties kunnen het instrument voor zelfevaluatie gebruiken om hieraan te werken.
- De zelfevaluatie van de organisatie wordt vervolgens door iemand extern aan de organisatie (de assessor) gevalideerd. De assessor gaat na of de organisatie daadwerkelijk beantwoordt aan de basiscapaciteit. Hij/zij stelt vervolgens enkele 'critical issues' vast op basis van zijn/haar analyse van de zelfevaluatie, het verbeterplan en bezoek aan de organisatie. De assessor gaat daarover in gesprek met de organisatie met de bedoeling het leren te ondersteunen. Deze aanpak laat toe om uit te gaan van de realiteit en context van de organisatie en haar autonomie inzake keuzes op het vlak van organisatie van de evaluatiefunctie te respecteren; dit schept het nodige vertrouwen om een oprechte analyse te doen en te werken aan verbetering.
- Gezien het belang van de kwaliteit van de evaluatierapporten, is het goed dat de assessor een tegenscoring doet van de kwaliteit van de evaluatierapporten, met de bedoeling om in dialoog te gaan met de organisatie over wat minder goed gaat, of over goede praktijken en de factoren die daartoe bijdragen.
- De aansturing en het beheer zijn zo *lean* mogelijk, want de stakeholders hebben weinig tijd om hierin te investeren: geen afzonderlijke structuren, aansturing voornamelijk door de DBE, samen leren door een jaarlijks leerevent rond evaluatie.
- Digitalisering: er wordt op dit moment niet gewerkt aan een concreet voorstel voor digitalisering hoewel dit interessant leek voor de uitvoering van het instrument en het

delen van ervaringen en goede praktijken. Er is op dit moment geen budget en er is de vrees dat digitalisering het instrument in deze fase te veel zou bepalen en zou bijdragen aan de rigiditeit ervan.

3. Bevindingen van de testfase

Tijdens de testfase werd het instrument uitgevoerd zoals beschreven in het document dat werd voorgelegd aan het begeleidingscomité van 8 september 2017. Dit document werd, mits enkele aanpassingen, vertaald in een handleiding voor zelfevaluatie voor de organisaties die meededen aan de testfase en in een format voor rapportage (combinatie van Excel en Word). Het instrument zelf wordt in dit hoofdstuk niet in detail toegelicht; we verwijzen hiervoor naar het volgende hoofdstuk en naar de bijlagen. Aan het einde van dit hoofdstuk wordt aangegeven welke aanpassingen werden doorgevoerd na de testfase.

De testfase werd georganiseerd tussen 12 september en 27 november 2017. De referentietermen van deze opdracht bepaalden dat de testorganisaties, bij een goede beoordeling, in aanmerking komen voor een certificaat dat de kwaliteit van hun evaluatiesysteem en -praktijk bevestigt.

3.1. Uitvoering van de testfase

De uitvoering van de testfase werd gekenmerkt door het volgende:

- De testfase stond zowel voor de actoren als voor de assessoren onder vrij grote tijdsdruk waardoor iedereen weinig tijd had om zich goed voor te bereiden maar iedereen heeft voldoende flexibiliteit betoond waardoor de analyse voldoende grondig gedaan werd.
- De assessoren konden vaststellen dat de zelfevaluatie ernstig en oprecht was gebeurd in de 8 organisaties en dat ze gebaseerd was op een collectief proces.
- De assessoren analyseerden in detail de methodologische kwaliteit van de evaluatierapporten die geselecteerd waren door de organisaties van de testfase. Voor BTC gebeurde de bepaling van de steekproef op basis van specifieke criteria die door de assessoren waren opgesteld.
- De aanpak was voor iedere organisatie dezelfde. Deze aanpak is beschreven in de handleiding van de assessor. Het bezoek aan de organisatie verliep doorgaans op dezelfde manier al was het startpunt soms verschillend (soms werd gestart met de analyse van de evaluatierapporten, soms met de huidige context van de organisatie). Er ging veel aandacht naar de dialoog met het team dat de zelfevaluatie uitvoerde, er was steeds een gesprek met de directie/het management, interviews met relevante personeelsleden die niet hadden deelgenomen aan de zelfevaluatie (indien mogelijk), en een debriefing waarin de eerste bevindingen gedeeld werden.
- De assessoren wisselden onderling uit over hun aanpak in de voorbereiding van hun bezoek.
- Verschillende rollen in de externe validatie werden getest: die van assessor, co-assessor en het finaal herlezen door de DBE (die garant stond voor een check van de coherentie en duidelijkheid van het verslag van de assessor).
- Vijf op de acht organisaties kregen telkens het bezoek van twee assessoren (Via Don Bosco, Entraide et Fraternité en Caritas niet).

- Tijdens het bezoek van de assessor namen vertegenwoordigers van de federaties deel als coach/ondersteuner van hun leden (behalve voor BTC en voor Avocats sans Frontières).
- De deelname van de DBE zorgde ervoor dat eventuele belangenconflicten vermeden werden (bijv. omdat de assessoren de organisatie kenden of betrokken waren geweest bij bepaalde evaluaties).

3.2. Conclusies over de elementen die getest moesten worden

Wat er te onthouden valt uit de evaluatie van de testfase over onderstaande aspecten die getest moesten worden:

Kwaliteit van het instrument voor zelfevaluatie: de handleiding bevatte veel informatie maar was onvoldoende helder gestructureerd en werd daardoor niet altijd even effectief gebruikt door de organisaties. Dit zorgde ervoor dat de zelfevaluatie niet altijd even sterk analytisch of goed gefundeerd was; door enkele organisaties werd ook te makkelijk verwezen naar het resultaat van de screening van de ANGS zonder dit verder toe te lichten. Het niveau van analyse verschilde sterk tussen de organisaties. Hieruit blijkt dat voldoende argumentatie en documentatie belangrijk blijft en dat organisaties ondersteund moeten worden in hun analyse. De uitleg over de evaluatiemethode kan nog duidelijker. Het format in Excel bleek onvoldoende gebruiksvriendelijk. Het werken met sub-criteria hielp om de criteria te analyseren. De organisaties vinden dat het format gerationaliseerd kan worden zodat ze minder moeten herhalen en dat het minder veeleisend wordt (bijv. vraag naar samenvattingen en globale analyse). Het was voor de organisaties onvoldoende duidelijk of en welke documenten zij moesten doorsturen.

Relevantie en helderheid van de criteria: de meeste criteria werden als relevant ervaren. De toelichting bij de sub-criteria was belangrijk om te begrijpen waar het over ging. Bij een minderheid van de criteria was er wel een verschil in interpretatie. De actoren gaven aan dat sommige criteria te binair waren opgesteld, waardoor het volstond met ja of nee te antwoorden (het is er of het is er niet) en er weinig bijkomende uitleg werd gegeven. De opdeling van de criteria in hoofdstukken en sub-hoofdstukken, werd als logisch en relevant ervaren. Er was een verschillende appreciatie voor het werken met de criteria die deel uitmaken van een basiscapaciteit inzake evaluatie en die een eerste drempel vormen voor het aanvragen van een bezoek van een assessor: sommige actoren vinden het nuttig, anderen niet. Organisaties vinden dat het antwoord 'niet-relevant' eventueel moet opgenomen worden in de scores, bijv. voor actoren die enkel een Noordwerking hebben (en die nu niet in de testfase zaten).

Haalbaarheid: uit de testfase leiden we af dat een organisatie minimaal 7 tot 8 werkdagen moet vrijmaken voor het totale proces (inclusief de analyse van de evaluatierapporten) indien de zelfevaluatie gebeurt door drie personen. Afhankelijk van de organisatie en haar keuze voor de opzet van het collectief proces loopt die tijd op, zeker voor grotere organisaties die de oefening zo breed mogelijk benut wordt in de organisatie en meerdere evaluatierapporten en -processen moeten analyseren. Een assessor heeft minimum 2,5 dagen nodig voor zijn voorbereiding, het bezoek en het verslag; voor meer complexe organisaties is meer tijd nodig. De assessor heeft minimum een halve dag nodig voor de analyse van 1 evaluatierapport (lectuur en rapportage). De testfase en de resultaten ervan doen vermoeden dat de vrijwillige testorganisaties extra gemotiveerd waren en ook een betere evaluatiepraktijk konden voorleggen. Dit maakt duidelijk dat een langere voorbereidingstijd voor sommige organisaties (waarin zij al een aantal verbeterpunten kunnen uitvoeren) en een

coaching van de actoren door de federaties (voor haar leden) of door externe consultants (bijv. voor BIO) wellicht geen overbodige luxe is.

Validatiemethode en interactie met de assessor: het idee van een dialoog tussen de organisatie en de assessor laat een rijke en open uitwisseling toe. Dat kwam ook omdat de assessor niet bleef hangen op ieder individueel criterium en de bijhorende scores (los van de basiscriteria) maar naar het systeem en de verbanden keek en belangrijke en kritieke punten (goede en minder goede) naar boven trachtte te halen (met wisselend succes). Het bleek doeltreffend om het gesprek met de organisaties te starten vanuit een toelichting bij hun context en hun Theory of Change of visie op verandering en daarbij aan te knopen met de concrete analyse van de evaluatierapporten. Er waren duidelijke spanningen in de dialoog tussen de assessor en de organisatie over leren vs. controle. De assessoren hebben dit op hun eigen manier beheerd, in één geval wel met minder succes. De assessoren hebben ervaren dat hun manier van valideren en scoren een stevige basis geeft voor een aanbeveling over de organisatie die rekening houdt met de kwaliteit van de zelfevaluatie, de beoordeling door de assessor, de kwaliteit van het verbeterplan en de inspanningen voor samenwerking. De organisaties geven aan dat de inspanningen voor samenwerking in de sector enkel in het verbeterplan aan bod kwamen maar te weinig verbonden waren met de criteria en dat er tijdens het bezoek door de assessor ook te weinig aandacht aan werd besteed. Bovendien vrezen zij dat dit aandachtspunt kan leiden tot artificiële constructies van samenwerking.

Meerwaarde van de oefening voor de organisaties: de organisaties gaven aan dat de oefening heeft gezorgd voor een grotere draagvlak voor de evaluatiefunctie binnen de organisatie, voor reflectie over evaluatie als een afzonderlijk deel van M&E, voor identificatie van specifieke verbeterpunten, voor interessante feedback over hun systeem door een externe assessor en voor het gevoel beter voorbereid te zijn voor de eerstvolgende evaluatie.

De testfase heeft volgende concrete outputs opgeleverd: zelfevaluatie en verbeterplan van de organisatie, verslag van de assessor per organisatie, aanbeveling van de assessor. De aanbeveling van de assessor en het verbeterplan van de organisatie kunnen gedeeld worden met de sector. Er werden ook een aantal goede praktijken geïdentificeerd die gebruikt kunnen worden.

De 8 testorganisaties kregen uiteindelijk hun certificaat voor ECB, Evaluation Capacity Building.

3.3. Gevolgen voor de uitwerking van het instrument

De testfase maakte volgende zaken duidelijk over de uitvoering:

- De testfase toonde aan dat de basisopzet van het instrument werkt en beantwoordt aan de vooropgezette doelstellingen. Het spanningsveld tussen het ondersteunen van leren en het 'ter verantwoording roepen' van de organisatie is inherent aan de kwalitatieve opzet van het instrument. Overigens zijn er grenzen aan wat het instrument kan bijdragen tot leren: de assessor is maar één van de externe factoren die eenmalig en gedurende een beperkte tijd uitwisselt met de organisatie. Er zijn ook grenzen aan het 'ter verantwoording roepen' of controleren, om dezelfde reden als hierboven en omdat de benadering sterk kwalitatief is.
- De testfase bevestigt dat het instrument werkt voor een diversiteit aan organisaties: door de zelfevaluatie situeren de organisaties de evaluatiefunctie binnen hun eigen context, waarmee de assessor vervolgens rekening kan houden.
- De testfase maakt duidelijk dat het niet mogelijk is om snel alle gevraagde informatie te krijgen van de overheid (bijv. documenten die werden opgestuurd door de

organisaties in het kader van de screening voor accreditatie), bestanden zijn vaak ook te zwaar om door te sturen.

- Een collectieve zelfevaluatie is een goede manier om te werken aan capaciteit voor evaluatie: het stimuleert uitwisseling en draagt bij tot een gedeeld begrip over wat evaluatie is. Een minimum van drie mensen die op verschillende niveaus werken in de organisatie blijft relevant. Het is echter best om de organisatie te laten bepalen hoe ze het collectieve leerproces organiseert.
- Het bepalen van een staal van evaluatierapporten: er werd verondersteld dat organisaties met veel evaluatierapporten een staal zouden nemen van het voorbije jaar. Dit biedt echter minder mogelijkheden om het gebruik van de evaluaties goed te analyseren. Het voorstel is dus om te kijken naar de laatste drie jaar.
- De handleiding legde de nadruk op het belang van het narratief en suggereerde om eerst het narratief in te vullen en nadien een consensus te vinden over een score. Het idee van scores bleek echter erg dominant en vaak werden de scores eerst toegekend en nadien onderbouwd. Dit draagt bij tot een verarming van het narratief, maar is moeilijk uit te sluiten.
- De testfase gaf aan dat een aantal criteria geschrapt konden worden.
- Het is belangrijk dat de assessor ter voorbereiding een aantal stappen op papier zet en die dan volgt. Die stappen zijn vermeld in de handleiding van de assessor. Ze moeten zeker behouden blijven.
- Er blijven vragen over de opvolging van de verbeterplannen. De sector liet duidelijk verstaan (in het begeleidingscomité) dat een jaarlijkse rapportage niet aan de orde is en dat een opvolging geïntegreerd kan worden in de jaarlijkse dialoog tussen de organisaties en DGD. De agenda van dit dialoog laat echter weinig ruimte voor een uitgebreide opvolging van het verbeterplan.

Na bespreking van deze resultaten met de organisaties van de testfase en met het begeleidingscomité van 7 november 2017, zijn de leden van het consortium het eens geworden over volgende noodzakelijke aanpassingen aan het instrument:

- Nog meer expliciteren dat evaluatie gaat over verantwoording afleggen en leren. Vandaar het belang dat evaluatierapporten in detail door de assessor geanalyseerd worden.
- De spanning tussen controle en ondersteunen van leren vertegenwoordigt een specifieke uitdaging voor de aanpak van de assessor, maar is ook inherent aan de aanpak: er moet daarom specifieke aandacht besteed worden aan vorming van assessoren.
- Omwille van budgettaire redenen, lijkt het ons niet aangewezen dat elk bezoek gebeurt door 2 assessoren. Dit kan wel voor meer complexe organisaties. Het is wel denkbaar dat een pool van vrijwillige peers (bijv. opgezet door en ondersteund vanuit de federaties) zich aandient op vraag van de actoren om deel te nemen aan het proces. Dit zal zeker het leereffect bijkomend ondersteunen. De eindaanbeveling blijft echter de verantwoordelijkheid van de assessor.
- De handleiding voor zelfevaluatie kan korter. Belangrijk is wel dat ze voldoende inhoudelijke ondersteuning blijft bieden (dus toelichting bij criteria en hulpvragen voor analyse). De richtlijnen voor het kiezen van een staal aan evaluatierapporten moet aangepast worden. Best is te werken met een combinatie van een Excel format voor de scores en een Word format voor het narratief. Er moet een duidelijk overzicht komen van documenten die geleverd moeten worden door de actoren.

- De lijst van criteria werd opnieuw bekeken alsook de toelichting. Bedoeling was vooral te vermijden dat simpele ja of neen antwoorden gegeven kunnen worden.
- De DBE en DGD bekeken samen welke gegevens over de actoren makkelijk gedeeld konden worden.
- De scores voor de zelfevaluatie: aanpassing van de schaal (van 1 tot 5 i.p.v. 0 tot 4, in te vullen met LEADS).
- Het komen tot de aanbeveling door de assessor: aangezien het waarderen van samenwerking rond evaluaties mogelijk tot artificiële constructies zou leiden, was het voorstel om dit te laten vallen in de waardering. Vanuit de overheid kwam ook een sterke vraag om de waardering van de assessor over het systeem sterker te laten doorwegen. Vandaar werd het voorstel van weging van scores het volgende: 20% voor de waardering van de zelfevaluatie (kwaliteit en authenticiteit), 50% voor de waardering van de assessor van de kwaliteit van het systeem en 30% voor de waardering van het verbeterplan (coherentie, relevantie en kwaliteit).
- Het instrument moet ook een formele component bevatten over uitwisseling en samen leren. We stellen voor dat de DBE jaarlijks een leerevent initieert. Dit kan georganiseerd worden met de sector en de federaties. Organisaties kunnen hier vrijwillig aan deelnemen: ze krijgen de kans om hun praktijken voor te stellen en toe te lichten hoe de uitvoering van het verbeterplan loopt en wat daarvan de resultaten zijn. De DBE kan de organisaties gericht uitnodigen om hun praktijk voor te stellen en dit samen voor te bereiden.

4. Beschrijving van het instrument

Achtereenvolgens komen in dit hoofdstuk de volgende punten aan bod:

- Situering van het instrument
- Doelstelling en scope
- Hoe zal het instrument werken?
- Uitvoering
- Modaliteiten voor beheer

4.1. Situering van het instrument

Naam van het instrument - De naam van het instrument is: ECB-initiatief van de Dienst Bijzondere Evaluatie. ECB staat voor Evaluation Capacity Building³⁴. Het instrument leidt tot het verlenen van een ECB-certificaat aan de actoren van de Belgische Ontwikkelingssamenwerking.

Het gaat om het toekennen van een certificaat door de Dienst Bijzondere Evaluatie aan een individuele organisatie. Het certificaat bevestigt de kwaliteit van het systeem van (externe) evaluatie, van de evaluatiepraktijk van een organisatie en van haar engagement tot verbetering.

Legale basis en context - Het initiatief en instrument passen binnen het kader van de Wet van 19/03/2013 betreffende de Belgische Ontwikkelingssamenwerking zoals gewijzigd door de wet van 9/01/2014 en de wet van 16/06/2016 en de nieuwe regelgeving met betrekking tot het mandaat van de Bijzondere Evaluator.

Uit het regelgevend kader volgt dat de actoren van de Belgische ontwikkelingssamenwerking verantwoordelijk zijn voor de evaluatie van hun interventies en dus kunnen aantonen dat ze hiervoor over de nodige capaciteit beschikken; en dat de Dienst Bijzondere Evaluatie (DBE) belast wordt met certificering van de evaluatiesystemen en het ex-post evalueren van de kwaliteit van die systemen.

Externe evaluatierapporten zijn een belangrijke bron van informatie voor de subsidiërende overheid. Sinds DGD niet beschikt over een interne evaluatiedienst, vormen de evaluatierapporten van de ANGS, BIO en BTC, in aanvulling op de evaluaties uitgevoerd door de DBE, de voornaamste informatiebron over de relevantie, doelgerichtheid, doeltreffendheid, impact en duurzaamheid van de interventies van de actoren. Dit onderlijnt het belang van de kwaliteit van de evaluatierapporten.

³⁴ Preskill & Boyle (2008), *A multidisciplinary model of evaluation capacity building*; Lennie (2015).

Het ECB-certificaat sluit aan bij de wens van alle betrokken partijen (DBE, DGD, NGO's, IA, BTC en BIO) om via monitoring en evaluatie een beter zicht te krijgen op de bijdrage van ontwikkelingsprogramma's aan veranderingen op outcome niveau. Sinds de meta-evaluatie van de DBE en de studie over evalueerbaarheid³⁵ zijn er in de sector al concrete stappen ondernomen om (de kwaliteit) van evaluaties sterker onder de aandacht te brengen en is er in toenemende mate overleg geweest tussen de partijen. Anderzijds investeren meerdere actoren in het ontwikkelen van nieuwe benaderingen die hen toelaten om beter zicht te krijgen op hoe en waarom interventies bijdragen tot verandering (aandacht voor veranderingstheorie, outcome mapping enz.). Er is dus een sterke motivatie van alle partijen om de kwaliteit van monitoring en evaluatie te verbeteren en zodoende te beschikken over betrouwbare gegevens om een uitspraak te doen over het behalen van resultaten, om kennis op te doen over een bepaald thema/onderwerp/interventie en om de implementatie van ontwikkelingsinterventies bij te sturen.

Aansturing van het instrument - Het initiatief voor de uitvoering van het instrument ligt bij de Dienst Bijzondere Evaluatie. Deze draagt de verantwoordelijkheid voor de uitvoering van het instrument en zijn resultaten binnen de diverse fora waar hij vertegenwoordigd is, meer bepaald bij DGD. Het gaat dan om de volgende fora, waar de DBE zowel de stand van zaken als resultaten van het instrument kan delen:

- Raad van Bestuur van BTC (individuele opvolging),
- Raad van Bestuur van BIO (individuele opvolging),
- NGSOC – overlegstructuur tussen de ANGS en de subsidiërende overheid (stand van zaken van het instrument en resultaten),
- Institutionele dialogen tussen DGD en de ANGS (individuele bespreking),
- Strategisch comité van DGD (stand van zaken, algemene lessen en/of trends),
- Parlement, in het kader van de voorstelling van het jaarverslag van de DBE (stand van zaken en algemene lessen en/of trends).

De DBE onderhoudt direct contact met de federaties, de Minister, DGD, BTC en BIO en nodigt hen op jaarlijkse basis uit op een event waaruit de actoren kunnen leren (zie verder onder "opvolging").

Om de aansturing en het overleg met de andere stakeholders te voeden met inzichten, maakt de DBE een transversale analyse van de uitvoering van het instrument en een synthese van belangrijkste bevindingen (inzake goede praktijken, trends in verbeterplannen, opvallende andere trends).

Benadering - Het instrument is gemodelleerd en geïnspireerd op diverse Europese en internationale voorbeelden van certificering en programma's voor de ondersteuning van de kwaliteit van externe evaluaties. Wat gedeeld wordt met bestaande voorbeelden is de aandacht voor een cyclus van continue verbetering (PDCA-logica of Plan, Develop, Check, Act). Het instrument onderscheidt zich van klassieke instrumenten voor certificering of labels waarbij criteria en instrumenten al bepaald zijn door internationale afspraken en het toekenningsmodel verloopt anders. Het instrument sluit eerder aan bij de benadering van het EFQM-model, hoewel het geen holistische benadering nastreeft en slechts een klein onderdeel bekijkt van de organisatie.

³⁵ DBE/ADE (2013) Meta-evaluatie van de programma's van de niet-gouvernementele actoren, DBE/South Research/IOB (2016) Geloofwaardig en zinvol evalueren; tussen droom en daad. Een studie van de evalueerbaarheid van interventies ge(co-)financierd door de Belgische Samenwerking.

Het instrument volgt de logica van Evaluation Capacity Building die gericht is op het duurzaam inbedden van een evaluatiepraktijk in de organisatie en het versterken van haar capaciteit om aan te sturen op betrouwbare, valide en bruikbare evaluatierapporten. Organisaties zijn steeds eigenaar van hun proces van capaciteitsversterking. Dit verloopt eerder niet lineair en zorgt niet altijd op korte termijn voor duidelijke resultaten. Een omgeving die het leren ondersteunt en leiderschap voor evaluatie zijn cruciale succesfactoren in capaciteitsversterking. Een collectieve aanpak vergroot het draagvlak voor capaciteitsversterking binnen de organisatie. Externe actoren kunnen het proces ondersteunen, oriënteren en valideren.

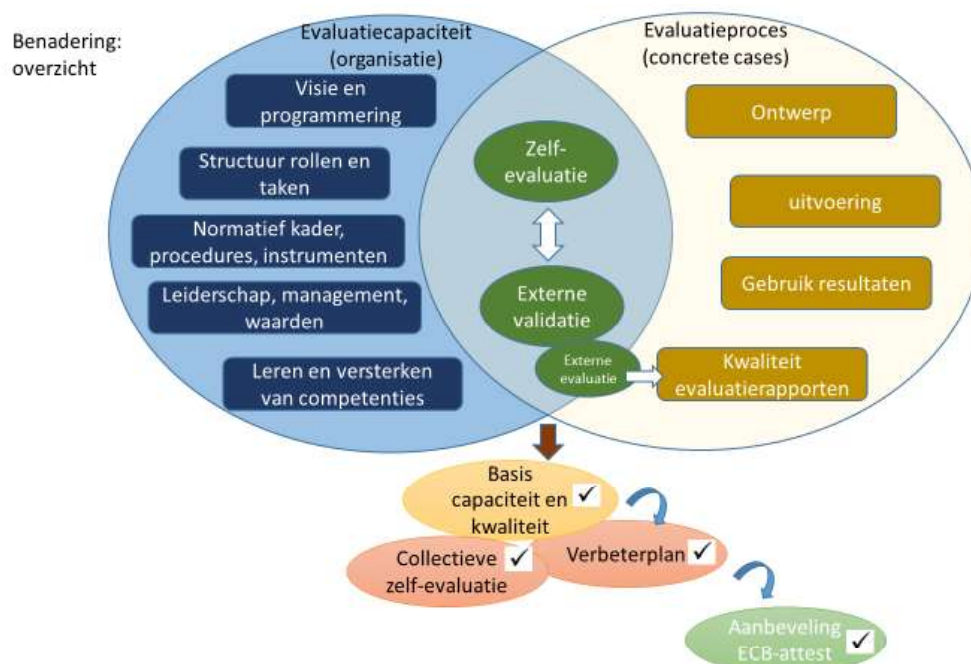


Fig.1: Overzicht aanpak

De ECB-logica vertaalt zich in het design van het instrument, de specifieke termen, de inhoud, de uitvoering en de opvolging. De verschillende etappes in het proces worden verder in het document toegelicht.

- **Design:** zelfevaluatie als het startpunt voor continue verbetering. Een zelfevaluatie door de individuele organisatie vormt de basis van het instrument en is het startpunt van een verbetertraject dat de organisatie zelf uittekent. De zelfevaluatie en het opstellen van het verbeterplan moeten gebaseerd zijn op een collectieve oefening. De organisatie brengt in haar analyse zelf bewijsmateriaal aan voor haar bevindingen en identificeert goede praktijken. Wel is het zo dat een organisatie minimaal moet beantwoorden aan een aantal basiscriteria (die zijn vastgelegd), dit vormt een garantie voor een zekere capaciteit van de organisatie inzake het opzetten, opvolgen en gebruiken van externe evaluaties.

- **Specifieke termen:** Een belangrijke term is 'assessor'³⁶; het gaat om een expert die in opdracht van de Dienst Bijzondere Evaluatie de zelfevaluatie en het verbeterplan van de organisatie analyseert en valideert en een gedetailleerde analyse uitvoert van de evaluatierapporten van de organisatie. Hij/zij gaat daarover in dialoog en discussie met de organisatie.
- **Inhoud:** uitgebreide set van criteria als kader voor de zelfevaluatie. Een lijst met criteria en toelichtingen biedt de organisaties inhoudelijke ondersteuning in het begrijpen waar het om gaat bij evaluaties. De criteria hebben specifieke aandacht voor leiderschap, het evaluatieproces, leren, ... Maar ook voor validiteit, betrouwbaarheid en bruikbaarheid van externe evaluaties. Een handleiding ondersteunt de organisatie in de zelfevaluatie en in de rapportage daarover.
- **Uitvoering:** externe validatie gebaseerd op bewijsvoering en analyse en identificatie van belangrijke domeinen voor verbetering. Het grootste deel van het werk ligt bij de organisatie zelf. Een tweede belangrijke rol is weggelegd voor een assessor. Een assessor valideert de resultaten van de zelfevaluatie en houdt de organisatie een spiegel voor op basis van de zelfevaluatie. Gezien het belang van betrouwbare en valide evaluatierapporten, beoordeelt de assessor in detail alle criteria die hiermee te maken hebben. De assessor vertrekt steeds vanuit de context van de organisatie en haar visie op verandering en probeert van daaruit en op basis van het beschikbare materiaal, 'critical issues' te definiëren en te bespreken. Om het leren te ondersteunen, geeft de assessor in zijn verslag voldoende toelichting (*appreciative feedback*).
- **Opvolging:** samen leren. Het delen van ervaringen en (goede) praktijken rond evaluatie en capaciteitsversterking met de andere actoren van de Belgische ontwikkelingssamenwerking is een belangrijke bijdrage tot ECB. De DBE zal hiertoe jaarlijks een initiatief nemen, in samenwerking met andere actoren.

De ECB-aanpak maakt geen gebruik van vragenlijsten rond criteria die binair geformuleerd zijn en enkel met ja of neen beantwoord kunnen worden. Bedoeling is niet 'compliance' te testen maar wel 'performance' van de organisatie te analyseren en te waarderen. ECB baseert zich op endogene processen en wat organisaties zelf aandragen als bewijsvoering en staat niet voor een aanpak waarbij de assessor enkel aantekeningen maakt. De onpartijdige analyse door de assessor en de interactie met de organisatie garanderen wel dat de assessor tot een oordeel kan komen over de kwaliteit en het engagement tot verbetering.

4.2. Doelstelling en scope van het instrument

Doelstelling - Het instrument heeft tot doel de kwaliteit van het systeem van (externe) evaluatie (capaciteit, beheer van evaluatieproces, aansturen op kwaliteit van evaluatierapporten), de praktijk van een organisatie en haar engagement tot verbetering vast te stellen en te bevestigen.

Het instrument heeft tevens tot doel, via de analyse van de evaluatierapporten door de assessoren, de vooruitgang van de kwaliteit (validiteit, bruikbaarheid en betrouwbaarheid) van evaluatierapporten op te volgen. Dit betreft de opdracht van de Dienst Bijzondere Evaluatie om ex-post de kwaliteit te evalueren.

De finaliteit van het ECB-certificaat is bij te dragen tot een continu leer- en verbeterproces bij de actoren en tot een betere kwaliteit van externe evaluatierapporten, een beter beheer van

³⁶ Deze term is afgeleid van de EFQM aanpak, omdat deze aanpak nauwer aansluit bij de ECB aanpak dan die van de klassieke certificering/audit aanpak.

evaluatieprocessen en een sterkere evaluatiecapaciteit. Dit kan op zijn beurt weer bijdragen tot meer impact op het terrein.

Scope/reikwijdte - Het instrument richt zich tot een specifieke doelgroep en een specifiek onderwerp.

- **Doelgroep:** de actoren van de Belgische ontwikkelingssamenwerking: ANGS, BTC en BIO. Precisering: voor de NGO's en IA's die deel uitmaken van grotere structuren die ontwikkelingssamenwerking NIET als core business hebben gaat het enkel om de dienst die zich met ontwikkelingssamenwerking bezighoudt.
- **Onderwerp:** het gaat om het M&E-systeem van organisaties, hun capaciteit, de processen (van ontwerp van een evaluatie tot gebruik van de resultaten ervan) en rapporten van externe evaluatie. Precisering: (i) het systeem voor monitoring wordt niet volledig meegenomen in het instrument, enkel de capaciteit van de organisatie om de gegevens te voorzien die nodig zijn voor de geplande evaluaties maakt deel uit van het instrument; (ii) het gaat om de processen en rapporten van externe evaluatie en meer bepaald de evaluaties van DGD-programma's. DGD heeft in overleg met de ANGS bepaald waar die rapporten aan moeten voldoen. Organisaties kunnen het instrument echter ook aanwenden om evaluaties van interventies gefinancierd door andere donoren te analyseren. Het dient benadrukt te worden dat het onderwerp deel uitmaakt van veel ruimere organisatiesystemen die op hun manier ook invloed hebben op evaluaties en de evaluatiepraktijk.

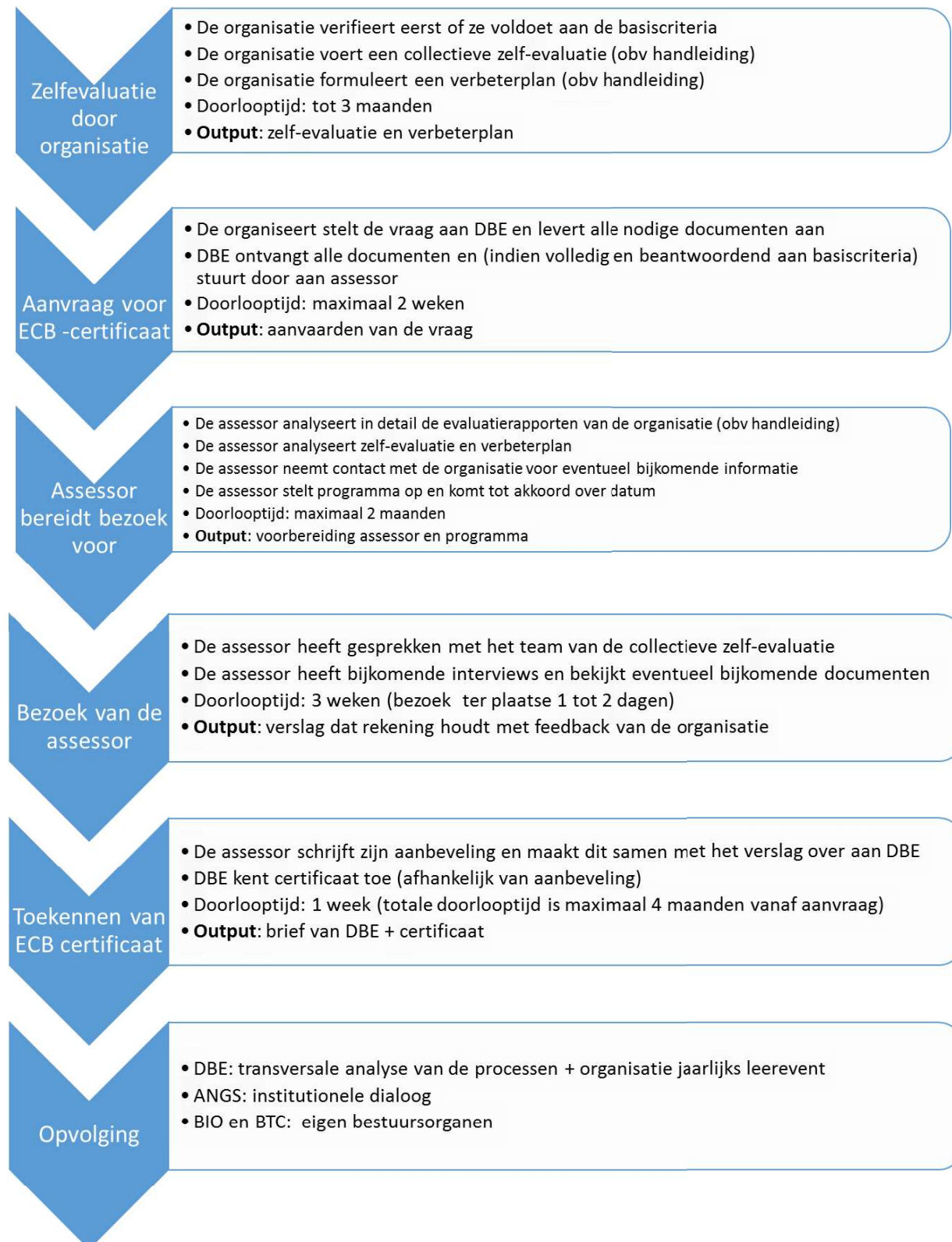
Ter verduidelijking van het onderwerp, verwijzen we naar de definitie van evaluatie volgens de OESO, zoals ook terug te vinden op de website van de DBE: *'Een evaluatie is een systematische en zo objectief mogelijke beoordeling van een project, een programma of een beleid dat aan de gang is of werd beëindigd, van haar concept, toepassing en resultaten, waarbij het de bedoeling is om de pertinentie, de realisatie van de doelstellingen, de efficiëntie inzake ontwikkeling, de doeltreffendheid, de impact en de duurzaamheid te bepalen. Een evaluatie dient geloofwaardige en nuttige informatie te verschaffen waarmee de ervaringslessen kunnen worden geïntegreerd in de besluitvormingsprocessen, zowel van de begunstigen als van de geldschieters. De drie functies van een evaluatie zijn (i) Verantwoording: rekenschap geven over het gebruik van overheidsfondsen, (ii) Leren: lessen trekken uit wat werkt en wat niet, (iii) Ondersteuning bij de besluitvorming: objectieve elementen rapporteren om de beslissingen te ondersteunen.'*

Incentives voor deelname – Diverse incentives kunnen een actieve deelname begunstigen:

- ECB-aanpak gericht op leren;
- Kwaliteit van de handleidingen en formats en oproep tot een *collectieve* analyse (die ertoe kan bijdragen dat de evaluatiefunctie goed wordt ingebed in de organisatie);
- Exposure: de organisatie kan haar aanpak en goede praktijken voorstellen op het jaarlijkse leerevent;
- Toetsen van de eigen visie aan die van een externe expert (uitdaging om uit de comfortzone te treden);
- Betere evaluatierapporten kunnen leiden tot meer inhoudelijke dialoog met de partners, waaronder DGD.

4.3. Hoe zal het instrument werken?

Onderstaand schema verduidelijkt de verschillende stappen (5 in totaal + opvolging) in het proces:



Criteria - Een lijst van criteria en sub-criteria ligt aan de basis van het instrument. De volledige lijst is terug te vinden in de handleiding voor zelfevaluatie. De inhoud van de criteria wordt in datzelfde document toegelicht. De criteria zijn opgedeeld in drie hoofdstukken: hoofdstuk 1 gaat over de evaluatierapporten van de organisatie, hoofdstuk 2 over de externe evaluatieprocessen en hoofdstuk 3 over de capaciteit van de organisatie om de evaluatiefunctie te organiseren. Onder hoofdstuk 3, capaciteit, zijn er 6 criteria die deel uitmaken van de minimale basiscapaciteit op het vlak van evaluatie. Deze criteria werden afgetoetst aan de praktijk van de organisaties maar ook aan de actuele literatuur³⁷ omtrent evaluatie.

Beoordeling en toekenningsmodel - Meer gedetailleerde informatie over de beoordeling en het toekenningsmodel wordt toegelicht in de handleiding voor zelfevaluatie en in de handleiding voor de assessor.

Er zijn **twee processen** die onafhankelijk van elkaar verlopen en een eigen dynamiek hebben: de collectieve zelfevaluatie door de organisatie en de validatie daarvan door een assessor.

Proces van zelfevaluatie - De criteria worden in eerste instantie beoordeeld door de organisatie in haar zelfevaluatie. Enerzijds door een narratieve argumentatie en vervolgens door het toekennen van een score op basis van een 5-punten schaal (van 1 tot 5, te begrijpen volgens de LEADS-beoordeling³⁸).

- L = Little or no elements can be identified
- E = Elements do exist
- A = concrete Action is taken
- D = largely Developed
- S = Sustainable

De narratieve beoordeling gebeurt aan de hand van concrete hulpvragen. Het instrument bepaalt niet welke de noodzakelijke bronnen van verificatie zijn of het bewijs dat de zelfevaluatie ondersteunt maar laat het aan de organisatie over om die aan te brengen.

De organisatie moet voor 4 op de 6 criteria van minimale basiscapaciteit een minimum score halen van 'D' om in aanmerking te kunnen komen voor de aanvraag voor het certificaat en het verkrijgen van het certificaat. Deze drempel moet ervoor zorgen dat organisaties zich goed voorbereiden voor ze een aanvraag indienen³⁹ en dat het certificaat een garantie biedt voor een minimale evaluatiecapaciteit.

Proces van externe validatie - De assessor analyseert op dezelfde manier (als de organisaties) maar onafhankelijk daarvan de criteria die betrekking hebben op de methodologische kwaliteit van de rapporten en besteedt bijzondere aandacht aan de validatie van de criteria van minimale basiscapaciteit. Het is verder niet de bedoeling dat de assessor ieder criterium afzonderlijk beoordeelt of dat hij/zij ieder criterium een eigen score geeft.

Om te komen tot zijn beoordeling en aanbeveling, analyseert de assessor de (i) kwaliteit van de collectieve zelfevaluatie, (ii) de kwaliteit van het systeem (in zijn geheel) en (iii) de

³⁷ We verwijzen hiervoor naar de beknopte bibliografie in bijlage.

³⁸ Dit wordt in de handleiding voor de zelfevaluatie uitgelegd met een concreet voorbeeld.

³⁹ Organisaties kunnen hiervoor eventueel beroep doen op advies en coaching van hun eigen federatie, peers of consultants.

kwaliteit van het verbeterplan. De beoordeling is zowel narratief als met een score volgens een 5-punten schaal (te begrijpen volgens de LEADS-beoordeling⁴⁰).

De scores worden omgezet in een totale score op 20 waarbij een verschillend gewicht wordt toegekend aan kwaliteit van de zelfevaluatie (20%), de beoordeling van de kwaliteit van het systeem (50%) en de kwaliteit van het verbeterplan (30%).

De assessor kiest uit een van de volgende drie mogelijkheden. Hij beveelt aan om:

- het certificaat toe te kennen (zonder voorbehoud);
- het certificaat toe te kennen met voorbehoud (dit wil zeggen dat de organisatie maximaal 3 maanden tijd krijgt om haar verbeterplan opnieuw te bekijken en sterker en/of coherenter te maken met de zelfevaluatie of om specifieke maatregelen te nemen);
- het certificaat niet toe te kennen (bijv. indien de assessor onvoldoende '*evidence*' vindt die aantoont dat de organisatie een minimale basiscapaciteit heeft).

Indien de organisatie geen minimale score haalt van 12/20, is er enkel een certificaat onder voorbehoud mogelijk, of kan beslist worden dat er geen certificaat kan toegekend worden.

Certificaten onder voorbehoud worden na maximaal 3 maanden (en indien er goed gevolg aan werd gegeven) omgezet in een certificaat zonder voorbehoud.

Afwijkingen - Gezien de opzet van het instrument, lijkt het ons niet nodig om afwijkingen toe te staan. In haar zelfevaluatie kan de organisatie altijd argumenteren waarom bepaalde criteria weinig of niet relevant zijn. Dit geldt echter niet voor de basiscriteria.

Beroep - Organisaties die een negatieve aanbeveling van een assessor willen aanvechten, sturen daarvoor een gemotiveerde brief naar de DBE. Indien de vraag gegrond is zal de DBE de procedure opnieuw opstarten met twee nieuwe assessoren.

Uitvoerders - De uitvoering van het instrument gebeurt door de DBE, een consultant en de actoren van de Belgische ontwikkelingssamenwerking (zelfevaluatie).

De DBE werkt hiervoor samen met een pool van minimum 6 assessoren.⁴¹ De assessoren moeten zowel competenties hebben op het vlak van evaluatie als op het vlak van organisatieontwikkeling (non-profit sector). Iedere expert beschikt minimaal over ervaring met evaluatie en met monitoring- en evaluatiesystemen gericht op het beheer van programma's voor ontwikkelingssamenwerking. Het personeel van de DBE kan in bepaalde gevallen tussenkomen als co-assessor.

De DBE staat open voor de deelname van peer-assessoren (experts uit peer-organisaties) die worden aangeworven door de actoren voor ontwikkelingssamenwerking en op hun vraag kunnen ingezet worden (afhankelijk van hun beschikbaarheid en hun eigen budget). Deze peer-assessoren zijn niet begrepen in het budget voor het instrument (zie verder). Bovendien ligt de verantwoordelijkheid voor het verslag en de aanbeveling enkel bij de assessor.

Uitvoering - De uitvoering per organisatie verloopt in verschillende stappen. De organisatie voert een zelfevaluatie uit aan de hand van criteria. Als blijkt uit die evaluatie dat de organisatie niet voldoet aan de criteria die een minimale basiscapaciteit vertegenwoordigen, werkt de organisatie op eigen initiatief en kracht aan verbetering zonder beroep te kunnen doen op een assessor.⁴² Indien de organisatie oordeelt dat zij klaar is voor een assessor, stuurt zij de zelfevaluatie met het verbeterplan, een formele aanvraag en bijhorende

⁴⁰ In de handleiding voor de assessor wordt dit uitgelegd met een concreet voorbeeld.

⁴¹ Dit aantal wordt bepaald door de nood om voldoende flexibel te zijn en het proces met 80 organisaties te kunnen managen. Het team mag niet te groot worden omdat het dan moeilijk wordt coherentie in benadering te verzekeren.

⁴² Eventueel kan de organisatie hiervoor terecht bij haar federatie, peers of bij consultants.

documenten naar de DBE. Het is belangrijk dat de organisatie voor het bezoek van de assessor al van start is gegaan met het uitwerken van haar verbeterplan, zodat de verbeterdynamiek duidelijk is. (Meer informatie is te vinden in de handleiding voor zelfevaluatie en de bijhorende formats.)

Eén of meerdere assessoren starten met de analyse van de documenten, gaan op bezoek in de organisatie en schrijven een voorlopig verslag, eventueel met aanbevelingen. De organisatie reageert op het verslag en voert eventueel aanpassingen uit aan haar verbeterplan. Op basis van deze aanpassingen finaliseren de assessoren hun verslag, waarbij ze rekening houden met de laatste versie van het verbeterplan en formuleren zij een aanbeveling aan de DBE. Het is de DBE die het certificaat formeel toekent.

Het aantal assessoren hangt af van de complexiteit van de organisatie. De lijst van complexiteit die werd opgesteld in het kader van de screening en het budget van de organisatie voor het DGD-programma is hiervoor richtinggevend.

Opvolging - Het is voorzien dat de actoren hun verbeterplan agenderen op de institutionele dialoog met de subsidiërende overheid. Voor BTC en BIO valt de opvolging automatisch onder de Raad van Bestuur, waarin ook de overheid vertegenwoordigd is.

Het instrument voorziet in een jaarlijkse vrijwillige collectieve opvolging van de verbeterplannen en een uitwisseling van ervaringen en evaluatiepraktijken door de organisaties. Dit zal gebeuren onder de vorm van een leerevent waarvoor de DBE uitnodigt (in overleg met de federaties, de ANGS, BTC en BIO) en waarvan telkens een verslag wordt gemaakt.

Kwaliteitscontrole op de uitvoering van het instrument - De kwaliteit van de uitvoering wordt op verschillende manieren ondersteund:

- Kwaliteitscontrole door de assessoren en de DBE;
- Startopleiding van 1 dag voor de experts/assessoren over de opzet van het instrument;
- Minimaal 2 x intervisie onder assessoren per proces (zie verder onder periodiciteit);
- Evaluatie van het instrument in 2020 door de DBE, de consultant en alle stakeholders: dit is essentieel om de effectiviteit van het instrument te verifiëren en om de relevantie ervan te garanderen (cf. evoluties in de sector en in het domein van evaluatie). Volgende vragen zijn belangrijk: geven de actoren aan dat de evaluatiefunctie beter georganiseerd is (zie 5 domeinen onder capaciteit)? Waar lagen de accenten van verbetering en wat is effectief gebleken? Zijn de evaluatieprocessen veranderd/verbeterd en sturen de actoren meer aan op kwaliteit van evaluatierapporten? Is de kwaliteit van de evaluatierapporten verbeterd? is er innovatie in methoden? Werd er geleerd van andere organisaties (wat en hoe)? Hoe verloopt het overleg met DGD (over referentietermen en institutionele dialoog)?

4.4. Modaliteiten voor het beheer

Transparantie en communicatie – De DBE kent het certificaat toe. Daarvoor steunt ze op de volgende informatie die ze aan de organisaties, DGD en de Strategische Cel communiceert:

- Aanbeveling van de assessor/organisatie: hierin argumenteert de assessor zijn aanbeveling en wijst hij op eventuele aandachtspunten;
- Verbeterplan van de organisatie (dat werd gevalideerd door de assessor): de organisatie maakt een globale analyse van haar sterke en zwakke punten, haar ambities binnen een bepaalde context en vertaalt dit naar een concreet verbeterplan;
- De scores en argumentatie van de assessoren ten aanzien van de methodologische kwaliteit van de evaluatierapporten (per organisatie en per rapport);
- De transversale analyse van de rapporten van de assessor met aandacht voor trends;
- De verslagen van het jaarlijkse leerevent georganiseerd door de DBE.

De DBE toont op de website hoe het afleveren van de certificaten vordert (namen van organisaties en jaar van certificering).

Periodiciteit – Een cyclus van certificering omvat vanaf 2022 in principe een **cyclus van 5 jaar** en loopt samen met de programmacyclus voor de ANGS en met het beheerscontract voor BTC en BIO:

- ANGS: de organisaties dienen een nieuwe aanvraag voor een ECB-certificaat in 2022⁴³ in nadat ze hun programma hebben ingediend. Het certificaat zelf maakt geen deel uit van de beoordeling van hun programma.
- BTC en BIO: het engagement voor continue verbetering van het evaluatiesysteem maakt deel uit van het beheerscontract. Dit is gegarandeerd voor BTC door de 'Fit for purpose' benadering en door de voorwaarden voor het nieuwe beheerscontract.

Organisaties moeten rekening houden met een termijn van minimaal enkele weken tot 3 maanden vooraleer zij de certificering aanvragen. Dit laat toe om alle voorbereidingen te doen. Eens de aanvraag ingediend is moeten zij binnen de 3 maanden het bezoek krijgen van een assessor.

Iedere cyclus vraagt opnieuw een volledige en collectieve zelfevaluatie en verbeterplan en een validatie door assessoren. In een tweede cyclus wordt er bijzondere aandacht besteed door de assessoren aan de verbetercyclus van de voorbije periode. Hiervoor zullen de handleidingen en formats aangepast moeten worden. Dat kan gebeuren na de uitvoering van de evaluatie van het instrument.

Van mei 2018 tot mei 2022 is er een **transitieperiode**. Tijdens deze transitieperiode gebeurt het volgende:

- Alle organisaties (ANGS en BIO) hebben ten laatste allemaal in mei 2019 een certificeringsaanvraag ingediend. Vanaf het moment dat een consultant door de DBE werd aangewezen (procedure van overheidsopdrachten), kunnen organisaties hun vraag indienen.

⁴³ De datum werd gekozen om rekening te houden met de programmacycli van de ANG's.

- In mei 2022 kan een tweede cyclus plaatsvinden (voor BIO en BTC in 2023, rekening houdend met hun beheerscontracten) die vanaf dan iedere vijf jaar hernieuwd wordt met het oog op het stimuleren van een continue verbeterproces bij organisaties.

Concreet verloop:

- Tussen januari 2018 en januari 2019: organisaties bereiden zich voor (eventueel met coaching van hun federaties of peers of consultants (voor BTC en BIO). Zij communiceren hun planning aan de DBE en vermelden wanneer ze een aanvraag voor certificering zullen indienen zodat de DBE hiermee rekening kan houden in de referentietermen voor de assessoren.
- Organisaties krijgen hun ECB certificaat (ten laatste in mei 2019 + uitstel voor organisaties die pas in mei hun aanvraag indienden): organisaties kunnen dus maximaal een jaar krijgen om zich voor te bereiden.
- Organisaties voeren hun verbeterplan uit en communiceren hierover in de institutionele dialoog met DGD.
- 2020: de DBE en de actoren evalueren samen de uitvoering van het instrument en formuleren verbeterpunten voor de certificering in 2022. De evaluatie gebeurt bij voorkeur samen met de ANGS en met BTC en BIO. Verbeterpunten voor het instrument worden afgesproken en geoperationaliseerd.
- 2021: organisaties bereiden hun nieuwe DGD/BTC/BIO-programma's voor en kunnen de criteria van de certificering meenemen in het achterhoofd (integratie van evaluatie in het nieuwe programma).
- 2022: een nieuw proces van certificering kan vanaf mei 2022 uitgevoerd worden.

Deze transitieperiode en de timing zijn niet onbelangrijk: het is belangrijk om de aandacht voor evaluatie gaande te houden na een eerste certificering, zeker omdat er geen formele jaarlijkse rapportage wordt verwacht. Na de transitieperiode kan gedacht worden aan een goede afstemming met het proces van de screening/accreditatie dat wellicht in 2025/26 van start zal gaan.

Afstemming met de afgelopen screening voor de accreditatie van ANGS door DGD -

De afstemming met de erkenningsprocedure is enkel van belang voor de ANGS. Uit de studie en de testfase blijkt dat de afstemming met de screening niet mogelijk is. Het gaat namelijk om twee verschillende processen. De screening gebeurt door DGD en gaat over een erkenning verbonden aan het recht om een subsidie te kunnen aanvragen. Zowel de totstandkoming als de aanpak en inhoud van de screening zijn anders dan die van het ECB-initiatief dat wordt uitgevoerd in opdracht van de Dienst Bijzondere Evaluatie (onafhankelijk van DGD) en enkel focust op monitoring- en evaluatiesystemen en dus niet op de gehele organisatie. Het ECB-instrument bouwt wel verder op enkele criteria van de screening en nodigt organisaties uit om ermee aan de slag te gaan en uit te leggen 'hoe' het *nu* gebeurt in de organisatie. Een eenvoudige verwijzing naar de resultaten van de screening van 2016 volstaat dus niet voor een grondig analyse in de zelfevaluatie.

Middelen voor de uitvoering - Voor de uitvoering van het instrument, zoals hierboven beschreven, moet rekening gehouden worden met minimum de volgende activiteiten: (een ruwe schatting brengt ons op een bedrag van 350.000 euro. Idealiter wordt iemand van de DBE voltijds vrijgesteld (1 persoon of verdeeld over de equipe).)

- Bestek opstellen en selectie van de consultants uitvoeren (DBE);
- Introductie evenement voor de actoren van de Belgische Ontwikkelingssamenwerking (DBE);
- Training van de assessoren;

- Ontvangen en verwerken van de aanvragen door de DBE;
- Proces per organisatie (3 dagen voor 1 assessor) voor 2/3 van de organisaties (complexiteit beneden 3,42 volgens de schaal van de screening/accreditatie);
- Proces per organisatie (4 dagen voor 1 assessor) voor 1/3 organisaties (met complexiteit boven 3,42);
- 2de assessor voor 1/3 organisatie (1,5 dagen voor voorbereiding, deelname aan bezoek en debriefing);
- Analyse van evaluatierapporten door assessor (gemiddeld 3 rapporten voor 41 organisaties met hogere complexiteit en/of budget van meer dan 8 ME voor de volledige programmaperiode, aan halve dag per rapport);
- Analyse van evaluatierapporten door assessor (maximaal 1 voor 47 organisaties aan halve dag per rapport);
- Kwaliteitscontrole door de assessor en de DBE;
- Transversale analyse inclusief rapportage (80 dossiers) ;
- Leerevents van 1 dag (1/jaar), inclusief rapportage;
- Evaluatie van het instrument (participatief moment + rapportage);
- Management van de documentatie (ontwerpen van een IT-tool?);
- Aanpassingen aan het instrument.

Succesfactoren – Een aantal succesfactoren zijn belangrijk voor de doeltreffendheid van het instrument.

- De DBE staat garant voor de kwaliteit van de uitvoering van het instrument en wordt daarin ook erkend door de stakeholders.
- Alle stakeholders (intern en extern aan de actoren) overtuigen dat evaluaties een goede zaak zijn: evaluaties die weinig nieuwe informatie brengen voor de organisatie kunnen meestal op weinig belangstelling rekenen binnen de organisatie en bij de partners. Organisaties die proactief aansturen op evaluaties met hun partners en zelf goed overwegen waarom en hoe hebben meer kans om bruikbare evaluaties te produceren. Vandaar het belang om bij de beoordeling van de capaciteit aandacht te besteden aan aansturing en oriëntatie van de evaluaties.
- Er is een minimum engagement nodig bij alle actoren: iedereen vindt het belangrijk maar de actoren staan onder druk. In een context van besparingen blijkt het in de praktijk moeilijk om middelen voor evaluatie opzij te zetten. Het uitvoeren van de zelfevaluatie zal een maximale inspanning lijken; organisaties zullen mogelijk niet zelf op zoek gaan naar bijkomende uitleg en zullen hun verbeterplan beperken tot acties die ze met het huidige personeel kunnen uitvoeren zonder bijkomende middelen. Dit kan een hindernis zijn voor innovatie. Daarom zijn een proactief aanbod van de federaties en het jaarlijkse leerevent, georganiseerd door de DBE, uiterst belangrijk.
- De federaties ontwikkelen een relevant aanbod ter ondersteuning van de actoren: het is op dit moment nog niet duidelijk waarover dit concreet kan gaan.
- DGD (D3 en D4) zijn transparant over de manier waarop zij gebruik zullen maken van de evaluatierapporten en hun recht om de referentietermen van evaluaties bij te sturen: een systematische aanpak werd nog niet ontwikkeld.

5. Overzicht van de bijbehorende documenten

In het voorgaande hoofdstuk werd beschreven hoe het instrument zal werken. Dit werd verder uitgewerkt in drie documenten: (i) overzicht van de criteria en toelichting, (ii) handleiding voor zelfevaluatie, (iii) handleiding voor de assessoren.

We beschrijven hieronder kort de inhoud van deze documenten die onderdeel uitmaken van het instrument.

5.1. Overzicht van de criteria

Het instrument is gebaseerd op een lijst criteria. Hieronder volgt een overzicht van de criteria en een precisering van de 6 criteria die een basiscapaciteit inzake evaluatie vertegenwoordigen. Deze criteria en de toelichting erbij ondersteunen de organisaties in hun zelfevaluatie en oriënteren hen in het verbeterplan. Hieronder het overzicht van de criteria, onderverdeeld in drie hoofdstukken.

Overzicht van de criteria	
Hoofdstuk 1: het evaluatierapport	
	<ul style="list-style-type: none">1.1. Validiteit: de validiteit van de informatie in het rapport is verzekerd.1.2. Betrouwbaarheid: de betrouwbaarheid van de informatie in het rapport is verzekerd.1.3. Bruikbare informatie: het evaluatierapport kan benut worden door de gebruiker.
Hoofdstuk 2: het evaluatieproces	
Ontwikkeling	<ul style="list-style-type: none">2.1. Kwaliteit van de referentietermen: de referentietermen beantwoorden aan minimale kwaliteitsstandaarden en oriënteren de evaluatie.2.2. Toezicht op de kwaliteit van de referentietermen: de organisatie staat borg voor een controle van de kwaliteit van de referentietermen.2.3. Participatie in het proces: de organisatie zorgt voor deelname van interne en externe partijen aan het proces van de evaluatie met het oog op onafhankelijkheid, kwaliteit en gebruik van de evaluatie.
Uitvoering	<ul style="list-style-type: none">2.4. Toezicht op de kwaliteit van de uitvoering: de organisatie staat borg voor een controle van de kwaliteit van de uitvoering van een evaluatie.
Gebruik	<ul style="list-style-type: none">2.5. Gebruik van de evaluatieresultaten: de organisatie gebruikt de evaluatie om haar eigen praktijk en die van andere stakeholders te verbeteren.
Hoofdstuk 3: de capaciteit om de evaluatiefunctie te organiseren.	

	<p>3.1. Oriëntatie: de organisatie expliciteert haar keuzes inzake evaluatie en het belang ervan voor de organisatie en oriënteert aldus de evaluatiefunctie.</p> <p>3.2. Structuur, rollen en taken: de organisatie structureert de evaluatiefunctie en definieert de rollen en taken van iedereen (de partnerorganisaties inbegrepen) om aldus de kwaliteit, coördinatie, onafhankelijkheid en het gebruik van evaluaties door de organisatie te garanderen.</p> <p>3.3. Normatief kader, procedures en instrumenten: de organisatie definieert en operationaliseert een normatief kader, verzekert de uitvoering en kwaliteit daarvan en ontwikkelt procedures en instrumenten om het werk met betrekking tot evaluaties te faciliteren.</p> <p>3.4. Leiderschap, beheerswijze en waarden: het leiderschap van de organisatie stimuleert de ontwikkeling van de evaluatiefunctie.</p> <p>3.5. Leren en versterken van competenties: de organisatie ondersteunt leren en versterkt de competenties inzake evaluatie.</p>
--	---

Er zijn zes sub-criteria in hoofdstuk 3 over de capaciteit van een organisatie die gedefinieerd werden als 'criteria die een basiscapaciteit vertegenwoordigen' (ze werden aangeduid met een blauw kader in het Excel format). Deze criteria hebben een grote invloed op de onafhankelijkheid, betrouwbaarheid en het gebruik van evaluaties.

Alvorens een organisatie de officiële vraag kan stellen aan de DBE om een certificaat voor ECB te ontvangen, moet de organisatie zich ervan verzekeren dat zij voor minstens vier van de zes criteria een goede score haalt, meer bepaald de score 'D' (die staat voor '*largely developed*' volgens het systeem LEADS).

Deze zes sub-criteria zijn de volgende:

- Sub-criterium 3.1.3: meerjarige programmering. Deze programmering vertaalt de visie van de organisatie en de manier waarop zij de functie van evaluatie articuleert en oriënteert. De organisatie definieert een evaluatieprogramma en actualiseert dat wanneer nodig en op basis van informatie uit het monitoringsysteem. Het programma verduidelijkt budgetten en verantwoordelijkheden alsook de keuzes in evaluatie (welke evaluatievragen, doelstellingen, focus, ...).
- Sub-criterium 3.2.1.: afstemming tussen evaluaties en besluitvorming: de organisatie zorgt ervoor dat het evaluatieprogramma afgestemd is met de besluitvormingscycli in de organisatie en de strategische en programmacycli. Op die manier zorgt ze ervoor dat resultaten van evaluaties effectief worden gebruikt bij het nemen van beslissingen
- Sub-criterium 3.2.2.: de organisatie definieert de rollen en taken van alle interne en externe stakeholders (de partnerorganisaties inbegrepen). De organisatie verzekert dat de taken worden uitgevoerd en garandeert op die manier de kwaliteit, een goede coördinatie en onafhankelijkheid in het beheer van de evaluaties.
- Sub-criterium 3.3.1.: de organisatie stelt de instrumenten, procedures en richtlijnen ter beschikking van alle betrokken stakeholders en waakt over de uitvoering en doeltreffendheid ervan. Het gaat om instrumenten, procedures en richtlijnen voor de voorbereiding, uitvoering en het effectief gebruik van evaluaties. Op die manier draagt de organisatie bij tot de kwaliteit van het evaluatieproces en de kwaliteit van evaluatierapporten.
- Sub-criterium 3.3.2.: de organisatie beschikt over een systeem voor opvolging dat in staat is om relevante, betrouwbare en bruikbare data ter beschikking te stellen van de

evaluaties die werden gepland. De organisatie draagt op die manier bij tot de kwaliteit van het evaluatieproces en de kwaliteit van evaluatierapporten.

- Sub-criterium 3.4.3.: personen die leidinggeven en/of betrokken zijn bij besluitvorming stimuleren de praktijk van evaluatie en het gebruik van de resultaten van evaluaties.

5.2. Handleiding voor zelfevaluatie

Dit document helpt de organisaties (ANGS, BTC/Enabel en BIO) in hun proces van zelfevaluatie van de kwaliteit van het evaluatiesysteem en in het formuleren van een verbeterplan. Het resultaat van de evaluatie en het verbeterplan zijn noodzakelijk om een aanvraag te kunnen indienen voor het ontvangen van een Evaluation Capacity Building (ECB) certificaat. De handleiding geeft aan wat de organisaties moeten evalueren, hoe ze hun evaluatierapporten en –praktijk kunnen analyseren en beoordelen, wat het belang is van een collectieve oefening binnen de organisatie en op welke manier ze hun verbeterplan kunnen opstellen. Er worden hiervoor ook formats aangeleverd, in Excel en in Word. De handleiding bevat tevens de volledige lijst met de criteria en de bijbehorende toelichting.

5.3. Handleiding voor de assessor

De handleiding voor de assessor verduidelijkt achtereenvolgens wat de taken zijn van de assessor, hoe hij/zij te werk moet gaan om de zelfevaluatie en het verbeterplan van de organisaties te analyseren, hoe hij/zij het bezoek voorbereidt en komt tot een finale beoordeling van de kwaliteit van de collectieve zelfevaluatie, de kwaliteit van het evaluatiesysteem en de coherentie en kwaliteit van het verbeterplan. De handleiding bevat een format voor het verslag van de assessor. Met dat verslag beveelt de assessor aan de DBE aan of de organisatie in aanmerking komt voor het verkrijgen van een certificaat van Evaluation Capacity Building.

Bijlage A: Overzicht van documenten die werden geconsulteerd

Wetten

Wet van 19/03/2013 betreffende de Belgische Ontwikkelingssamenwerking, zoals gewijzigd door de Wet van 9/01/2014 en de wet van 16/06/2016 ;

KB van 25/02/2010 houdende de oprichting van een Dienst Bijzondere Evaluatie van de Ontwikkelingssamenwerking, zoals gewijzigd door het KB van 25/04/2014, met inbegrip van het Verslag aan de Koning ;

KB van 11/09/2016 betreffende de niet-gouvernementele samenwerking;

KB van 10/04/ 2014 houdende instemming met het Vierde Beheerscontract tussen de Belgische Staat en de naamloze vennootschap van publiek recht met sociaal oogmerk Belgische Technische Coöperatie alsook Vierde Beheerscontract tussen de Belgische Staat en BTC van 9/04/2014 ;

KB van 2/04/2014 houdende instemming met het Eerste Beheerscontract tussen de Belgische Staat en de naamloze vennootschap van publiek recht "Belgische Investeringsmaatschappij voor Ontwikkelingslanden" alsook Eerste Beheerscontract tussen de Belgische Staat en BIO van 1/04/2014 ; zoals gewijzigd door een bijkomende overeenkomst van 20/12/2016;

Andere documenten

ADE (2016) *What does the future hold for impact evaluation in the Belgian NGO sector? The lessons from four case studies*. Synthesis report, Volume 1.

Better evaluation website (www.betterevaluation.org): using the Rainbow Framework Blog post by Irene Guijt ([weblink](#)).

Bockelie, J; Boisteau, C; Pioch, L. (2017) *Comment renforcer les apprentissages et les changements à travers l'évaluation? Analyse de pratique pour une évaluation transformative*. Les collections du F3E.

De Bruyn, T.; Huyse, H. (2015) *New Trends in Governmental Funding of Civil Society Organisations: A Comparative Study of 9 OECD-DAC Donors*. HIVA-KULEUVEN ([weblink](#))

DBE (2017) *Discussienota van de DBE over de certificering van de M&E systemen van de Belgische Ontwikkelingsactoren*.

DBE/ADE (2013) *Meta-evaluatie van de programma's van de niet-gouvernementele actoren*.

DBE/South Research/IOB (2016) *Geloofwaardig en zinvol evalueren; tussen droom en daad. Een studie van de evalueerbaarheid van interventies ge(co-)financierd door de Belgische Samenwerking*.

Deprez, S. (2016) *Creating the (organisational) conditions for an OM-based M&E and learning practice*. (Outcome mapping learning community).

DFID (2014) *Rapid Review of Embedding Evaluation in UK Department for International Development*. ([weblink](#))

DGD, D4 (goedgekeurd door de minister op 23/04/2015) *Strategienota ontwikkelingsresultaten. DG Ontwikkelingssamenwerking en humanitaire hulp*.

DGD, D3 (2016) *Screening des acteurs ANG's 2016: vue global des scores*

EBA (2017) - *Confronting the Contradiction – An exploration into the dual purpose of accountability and learning in aid evaluation*. Reinertsen, H., Bjørkdahl K., McNeill, D.. - Expertgruppen för bistandsanalys (EBA).

Inspectie Ontwikkelingssamenwerking en beleidsevaluatie (S.d.) *Methodische kwaliteit van programma-evaluaties in het medefinancieringsstelsel - I 2007-2010*.

Hallam, A. & Bonino F. (2013) *USING EVALUATION FOR A CHANGE: Insights from humanitarian practitioners*. ALNAP ([weblink](#))

IDRC/CRDI (International Development research Centre): [Evaluation at IDRC](#), [Evaluation Guidelines / Bulletins Ligne directrice pour l'évaluation](#)

ILO (2014) *Checklist to rate quality of evaluation reports* ([weblink](#))

ITAD (2014) *Investing in monitoring, evaluation and learning - Issues for NGO's to consider*.

June Lennie, Jo Tacchi, Michael Wilmore, Bikash Koirala (2015) *A holistic, learning-centred approach to building evaluation capacity in development organizations*. Evaluation. Vol.21, Issue 3, pp 325-343 ([weblink](#))

Kabinet van de Minister voor Ontwikkelingssamenwerking De Croo (S.d.) *Overkoepelende nota Beleidsceel Ontwikkelingssamenwerking, "Trust, but verify". Evolutie van de toezichthoudende aanpak en structuren binnen OS*.

Minister voor Ontwikkelingssamenwerking, De Croo (17 oktober 2017). *Algemene Beleidsnota Internationale ontwikkeling*. DOC 54 2708/003.

OECD (2010) *Lignes directrices et ouvrages de référence du CAD. Normes de qualité pour l'évaluation du développement*.

OECD (2016) *Evaluation Systems in Development Cooperation* (2016 Review) ([Weblink](#))

OCDE-CAD (2002) *Glossaire des principaux termes relatifs à l'évaluation et à la gestion axée sur les résultats*.

OCDE-CAD (2010) *Evaluer la coopération pour le développement*.

Peersman, Greet; Guijt, Irene & Pasanen, Tiina (2015) *Evaluability assessment for impact evaluation. Guidance, checklists and decision support*. Methods Lab. ([weblink](#))

Preskill & Boyle (2008), *A multidisciplinary model of evaluation capacity building*

Service de l'évaluation spéciale (mai 2014) *Politique d'évaluation. Tirer des enseignements de l'expérience passée et rendre compte des résultats*.

UNEG, Davies, I. & Brümmer, J. (2015) *Final report to the UNEG Working Group on Professionalisation of Evaluation*. ([weblink](#))

UNEG (June 2015) *Evaluation competency framework*

UNEG (2016) *Normes et règles d'évaluation*.

Over BIO

BIO Annual Report 2015

BIO Annual Report 2014

BIO Development Assessment (Direct investments, Financial Institutions, Funds and Infrastructure)

BIO Project Completion (ibid)

BIO, Procedure for project completion reporting', February 3, 2017

BIO, Procedure for Project Completion Reporting, March 2016

BIO, BIO's Interne evaluatiesysteem, December 2015

BIO, Outline of a development results framework for BIO, sept 2014, 12p

BIO – Case study evaluation of 5 BIO investments in Senegal TOR – Oct 2014

BIO – Case study evaluation of 5 BIO investments in Senegal -Note on Sample Selection, Oct 2014

SPF Affaires étrangères, Commerce International et Coopération au Développement Internationales, DGD, D2.2 *Croissance inclusive Note à monsieur Alexander DE CROO, Vice-Premier Ministre et Ministre de la Coopération au développement, de l'Agenda numérique, des Télécommunications et de la Poste* ref D2.2/GJ/DEV/10.01/2016 02 501 43 71 Objet: Société belge d'Investissement pour les Pays en Développement (BIO) – Octroi des moyens financiers de code 8 prévus pour 2016 par le Contrat de gestion (40 M euros – AB 145444 81.61.43).

Impact Consulting sprl (2015) *Mission d'évaluation des investissements de BIO au Sénégal, rapport final.*

PHB Development (2016) *Case Study Evaluation of Five BIO Investments in East Africa*

Enclude (2017) *Draft Final Report, Case Study Evaluation of five BIO Investments in Agriculture - Volume 1 – Synthesis Report, Volume 2 – Case Study Reports* Executive Summary

DFIs should work together to measure job. <http://blogs.worldbank.org/jobs/dfis-should-work-together-measure-job-impacts>

Let's Work Global Partnership, letswork@ifc.org (BIO is not a member)

Massa I, Mendez-Parra M & te Velde D.W. (2016) *The macroeconomic effects of development finance institutions in sub-Saharan Africa Report.*

Overseas Development Institute (Alberto F. Lemma) (2015) *Development Impact of DFIs What are their impacts and how are they measured?*

Project directors Daniel F. Runde F.; te Velde D.W, Principal authors, Conor M. Paddy Carter S Lemma A., (2016) *Development Finance Institutions Come of Age; Policy Engagement, Impact, and New Directions - A Report of the CSIS project on Prosperity and Development and the Overseas Development.*

Over BTC

BTC (2016) *MORE RESULTS – Monitoring and Reviewing Results : A Guide.*

BTC (2017) *Evaluatiecharter van de Interne Evaluatiedienst van BTC*

SES (South Research, Ecorys) (2010) *Évaluation de la qualité des prestations de la Coopération Technique Belge.*

SES (Particip, ADE) (2012) *Evaluation de la qualité des prestations de la Coopération Technique Belge* .

SES (2013), *Evaluer l'impact, la quête du Graal ? –Evaluation ex post de l'impact de quatre projets de coopération gouvernementale* –Rapport de synthèse, SPF Affaires étrangères, Commerce extérieur et Coopération au développement, ADE, Bruxelles, Janvier 2014

SES (2016) *Evaluation ex-post de la durabilité de la coopération gouvernementale, Que se passera-t-il une fois notre coopération au développement terminée ?* Service Public Fédéral Affaires étrangères, Commerce extérieur et Coopération au Développement, Bruxelles., HERA – ADE, Février 2016

Bijlage B: Overzicht van organisaties die werden geconsulteerd

Datum	Etappe	Organisatie
3 mei 2017	dataverzameling	DGD (D4)
4 mei 2017	dataverzameling	DGD (D3)
8 mei 2017	dataverzameling	NGO-Federatie
9 mei 2017	dataverzameling	DGD (D3)
9 mei 2017	dataverzameling	ACODEV
10 mei 2017	dataverzameling	Kabinet van Minister De Croo
10 mei 2017	dataverzameling	FIABEL
11 mei 2017	dataverzameling	DGD (D2)
15 mei 2017	dataverzameling	DGD (D1)
15 mei 2017	dataverzameling	BTC
17 juni 2017	dataverzameling	BIO
21 juni 2017	dataverzameling	DGD, adjunct Directeur - Generaal
6 juni 2017	Dataverzameling ontwerp en	FIABEL, ACODEV en NGO-federatie
6 juni 2017	Dataverzameling ontwerp en	DGD, dossierbeheerders van D3
19 juni 2017	Dataverzameling ontwerp en	DGD, D4
19 juni 2017	Dataverzameling ontwerp en	BIO
28 juni 2017	Dataverzameling ontwerp en	SES, BTC, Kabinet, DGD, D4

Bijlage C: Overzicht van key-events

Informatiesessie		
DATUM	Etappe	Activiteit
27 april 2017	Opstart van de studie	informatiesessie

Focus Group		
DATUM	Etappe	Activiteit
17 mei 2017	dataverzameling	Focus group n°1
23 mei 2017	dataverzameling	Focus group n°2
24 mei 2017	dataverzameling	Focus group n°3
30 mei 2017 VM	dataverzameling	Focus group n°4
30 mei 2017 NM	dataverzameling	Focus group n°5

Sensemaking Workshop		
DATUM	Etappe	Activiteit
21 juni 2017	Verzamelen informatie van	Sensemaking Workshop 2
22 juni 2017	Verzamelen informatie van	Sensemaking Workshop 2
Restituties		
DATUM	Etappe	Activiteit
27 april 2017	restitutie	Restitutie met DGD, D3
19 januari 2018	restitutie	Restitutie met sector en alle actoren

Bijlage D: Overzicht van organisaties uit de testfase

Naam	# evaluatie- rapporten door assessor geanalyseerd	aantal	Datum zetelbezoek	van	Aanwezigheid federatie
Entraide et Fraternité	2		16/10/2017		Ja
VLIR- UOS	3		19/okt/17		Ja
BIS-MSI	2		17/okt/17		Neen
Avocats sans Frontières	1		12/okt/17		Neen
Artsen zonder Vakantie	1		11/okt/17		Ja
Via Don Bosco	1		12/10/2017 23/10/2017	+	Ja
Caritas Internationaal	3		23/10/2017		Ja
BTC	5		9/10/2017 18/10/2017	+	Neen